The Pennsylvania State University

The Graduate School

The Department of Civil Engineering

INVESTIGATION OF COMPUTER VISION METHODS FOR THE BUILDING CONSTRUCTION PROCESS

A Thesis in

Civil Engineering

by

Donald J. Wallace

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 1988

We approve the thesis of Donald J. Wallace.

Date of Signature

ABSTRACT

The objective of this research was to investigate the feasibility of introducing computer vision methods to the building construction site. The primary use envisioned is direct input of data from a video medium to the computer for the purpose of productivity analysis.

The first step described is a familiarization with a fundamental computer vision system.

Then, the results of processing actual footage from construction sites and other, existing buildings are described. There are observations made as to the effect of backlighting of the structure, lighting conditions and shadowing, and physical obstructions. These concepts are illustrated by digitized images of the structures observed.

Proposals are then made as to methods to insure precise repeatable placement of observation cameras. The alternate proposal is translation of images obtained from different camera positions through the use of on-screen reference points.

Finally, a summary of the physical barriers and technological problems, and suggested courses of action, is provided.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION AND BACKGROUND

Traditional Data Acquisition on the Construction Site

Productivity can be defined as the ratio of input resource units to output units (input/output). In the context of building construction, this can be specified as manhours-, machine hours-, or dollars-per-ton, -lineal foot, -square foot, or -cubic yard.

Productivity on the construction site is currently determined by several means. The simplest, yet least useful for productivity improvement programs, is after-the-fact comparison of total resources expended versus quantities in place. Several more useful methods involve data collection, on the job site, of work in progress. This data could be in the form of: Time-lapse photography; Work Sampling; or Manual timing of operations.

The method of time-lapse photography involves the viewing of video footage (or film) of construction operations. Observations are then made as to inefficiencies in procedures, wasted travel time to locations of work and materials, and the actual production rate of work in place.

Work sampling involves random surveys of the job site to observe whether "effective work" is underway. Effective

work is defined as those activities that either actually place materials in the structure or directly support such activities. Ineffective work is everything else (standing around, drinking a soda, or transporting materials long distances). By gathering sufficient data on all activities, a picture of the effectiveness of the work force can be drawn.

Manual timing of operations involves observers on the job site with stopwatches and notepads recording actual performance times of the activities of concern. This method is extremely manpower intensive.

These productivity analysis methods all involve some statistical analysis of the job. Statistical sampling is necessary due to the large expenditure of time and capital involved in continuous observation. This large expenditure occurs not only during the collection of the data, but also during the required analysis.

The collection of data is the first of four steps in the statistical sampling process. The remaining steps are: (2) organization of data, (3) analysis of data (determination of "descriptive statistics"), and (4) interpretation of data ("inferential statistics") (Dillman 1981, pp. 7-8). Theoretically, proper evaluation of a statistical sample provides reasonably accurate results.

The final three steps listed above are well suited for accomplishment by computer. However, manual data collection

and manual input to the computer fail to utilize the full potential of the technologies being developed in the manufacturing industry.

Historical Development/System Trends in Computer Vision

"Pattern recognition" can be defined as the computerized process of analyzing a sensory image and differentiating and recognizing the component parts of that image. This concept is widely used in the factory environment for quality control and for the control of robotics.

The current state-of-the-art pattern recognition system requires that very distinct patterns be provided. In the factory this is accomplished by separating individual items and closely scrutinizing them under optimal lighting conditions. This is impractical on the construction site, where lighting is largely uncontrollable and the distances to the object are greater; but then the requirements at this stage are not to discern small individual components but to survey bulk quantities of production, e.g. square feet of wall, feet of pavement or pipe, etc. These differences in objectives and physical characteristics necessitate a different class of optics and recognition systems than those used in the factory environment.

Objectives of Study

The objective of this research was to investigate the feasibility of introducing computer vision methods to the building construction site. This objective was identified as a primary consideration for introducing automated data collection for productivity analysis of the construction process (Thomas and Smith 1987).

The tasks necessary to fulfill this objective are:

1. Learn the capabilities of an entry-level computer vision system with regard to its ability to reconstruct a given image with varying levels of contrast.
2. Select sample structure(s) for data collection.
3. Determine the feasible angles of view for observation cameras, and their related distances, based upon the nature of the object observed and the specifications of the camera system.
4. List the steps involved in correlating the dimensions of the visual image processed by the computer vision system with the actual dimensions of the observed object, and the corresponding areas.
5. Identify the physical and technological barriers to effective use of entry level vision systems hardware in construction.

Methods

The first requirement was an in-depth literature search and review. Publications on computer vision and construction management were reviewed for information on previous work on vision systems employed for productivity data collection. In addition to published literature, course notes from the Computer Vision and Inspection course offered at the Pennsylvania State University were reviewed as a primer on vision system techniques and applications.

In conjunction with the literature search, the laboratory exercises for the aforementioned computer vision course were performed to familiarize the author with the equipment.

Field data collection was performed using a Panasonic video camera and recorder. Time of day, temperature, weather, and location were all considered prior to collection of data. The following structures were considered as potential subjects: Mid-State Bank (under construction); Atherton Hotel (under construction); and Centre Community Hospital (not under construction). Upon selection of Centre Community Hospital for analysis, as-built drawings were obtained for additional information.

Analysis was comprised of executing pixel counts on the images collected. These images were then qualitatively evaluated for indications of where system improvements are

required, and where potential technological limitations exist.

The findings of the study enabled the author to determine that, although feasible, substantial work remains prior to prototype applications.

CHAPTER 2

HARDWARE AND SOFTWARE

Determination of PCEYE System Capabilities

The computer vision system utilized for this study is PCEYE, marketed by Chorus Data Systems. Its unit cost is under $1000, exclusive of the computer and camera equipment. The PCEYE system used for this study, as available in the Computer Vision and Inspection Laboratory (CVIL) of the Mechanical Engineering Department of The Pennsylvania State University, includes the following components:

- -IBM Personal Computer with 256K RAM, dual 360K drives
- -Color Monitor, using Color Graphics Adapter
- -PCEYE System Board
- -Black and White Video Camera
- -Black and White Video Monitor
- -Video Cassette Recorder (VCR)
- -Epson Dot Matrix Printer

All laboratory work (e.g. image processing) was accomplished in CVIL using this equipment.

Input Characteristics

The PCEYE system accepts as input the standard analog RS-170 signal, in common use for black and white video transmission, or the National Television Systems Committee (NTSC) signal, which carries the information for color

images in addition to the RS-170 data. The NTSC signal is common to television and VCR transmissions. The RS-170 signal is illustrated in Figure 2.1. This signal causes the scanning beam in a television picture tube to scan across the tube in consecutive horizontal lines from top to bottom. It does this thirty times each second.

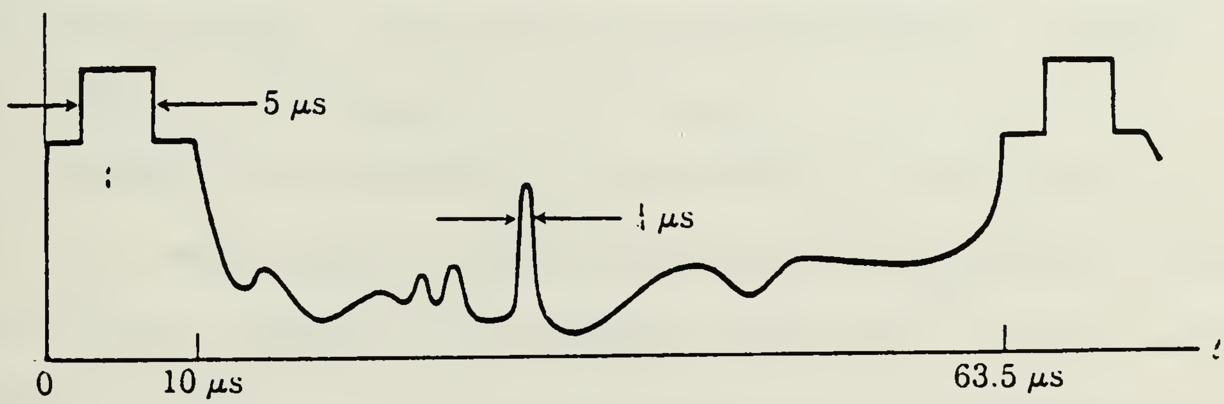
This continuous function (signal) is translated into a discrete (digitized) function by an analog-to-digital converter on the PCEYE system board.

The PCEYE system is also capable of retrieving images which were previously saved to disk. These disk-saved images are currently what must be used for analysis and comparison.

The laboratory setup used for course work typically feeds the signal directly from the black and white camera through the video monitor to the PCEYE board. For the purposes of this study, it was necessary, and in fact desirable, to record the images on videotape at the building site, and replay them on a VCR in the laboratory for input to the system. This was desirable because it more precisely matches the probable procedure that will be used in construction practice.

Resolution

Fineness(Size,Crispness). The PCEYE system, as configured, breaks the image into 320 picture elements



RS-170 SIGNAL (Carlson and Gisser, 1981)
FIGURE 2.1

(pixels) horizontally by 200 pixels vertically. Using the example of a building 100 feet tall which fills the screen vertically, each vertical pixel then equates to six inches.

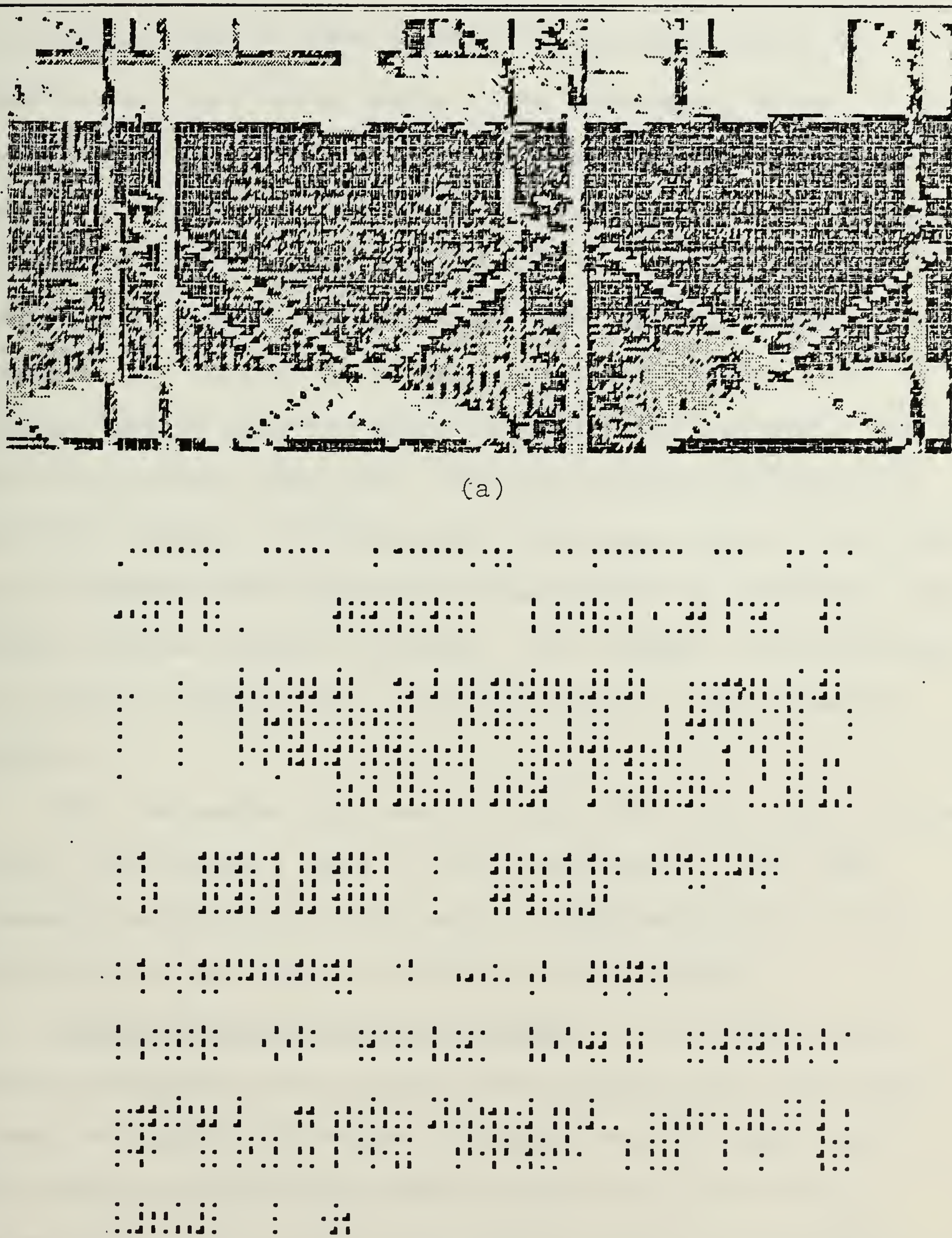
The resolution of the 320 by 200 system is the minimum recommended for use. Figure 2.2(a) is the upper left 64 by 64 image extracted from the PCEYE image in Figure 2.2(b).

Intensity. The existing system uses four grey levels to represent a digitized image. These four grey levels, or light intensities, are displayed as four pseudo colors on the color monitor. In ascending order they are: black, blue (cyan), red (magenta), and white.

The user can adjust the sensitivity of the digitizer until the image observed meets the desired intensity levels, the resultant image could be black and white, or four color, whichever brings out the desired details. The sensitivity is actually assigned by directing the lowest intensity that is interpreted as white, and the highest intensity that is interpreted as black. The system then distributes the other two grey levels between these two limits.

Image Processing

Color versus Black and White. There is usually no problem inputing a color signal to a simple frame grabber such as PCEYE, since the color signal usually contains the same information in the same format as a black and white signal. Any problem that arises is based upon the limi-



(a)

(b)

COMPARISON OF (a) PCEYE IMAGE AND (b) 64 x 64 EXTRACTION
FIGURE 2.2

tation that such a frame grabber is designed solely to handle black and white images. The extraneous color information can be misinterpreted, causing interference and/or a fuzzying of the image; therefore the color portion of the signal is normally filtered out. Systems that make use of the color signal are under development. What will be found in the initial development of color systems is to a greater degree what was found in grey level systems. The more grey levels there are, the more complex and slow the software becomes. The distinct advantages arise in that the system becomes more tolerant to variations in lighting, thus requiring less special lighting. For example, backlighting is usually required for a black and white, edge detecting system.

For the system utilized in this study, the color of the object only plays a role in the development of an image inasmuch as the color can create a differentiation in the intensity of the light reflected by that object.

Volume of individual grey levels. The PCEYE system package contains two programs which process the digitized image and report the number of pixels at each grey level. The images processed by these programs must have been previously saved to disk in a format different from the standard PCEYE format (see CAPTURE.BAS below).

The first program processes the entire image (frame), which takes two to three minutes. The product is a histo-

gram showing graphically the number of pixels per grey level.

The second program will process any rectangular portion of the image designated by the user, up to the full screen. This can take over an hour, but more detailed information is provided. The user designates what "colors" are part of the object, and is told the size of the object, its centroid, and if there are any "holes" in it (visually). This program was developed at the Pennsylvania State University and is not available with PCEYE.

Image Enhancement. Using this system, the only practical method of enhancing the image is proper adjustment of the black and white levels. These levels determine the upper limit of light intensity that will be interpreted as black and the lower limit interpreted as white. The user sets the limits by either entering the levels into the parameter table available in the PCEYE program, or by adjusting the image as it appears on the screen.

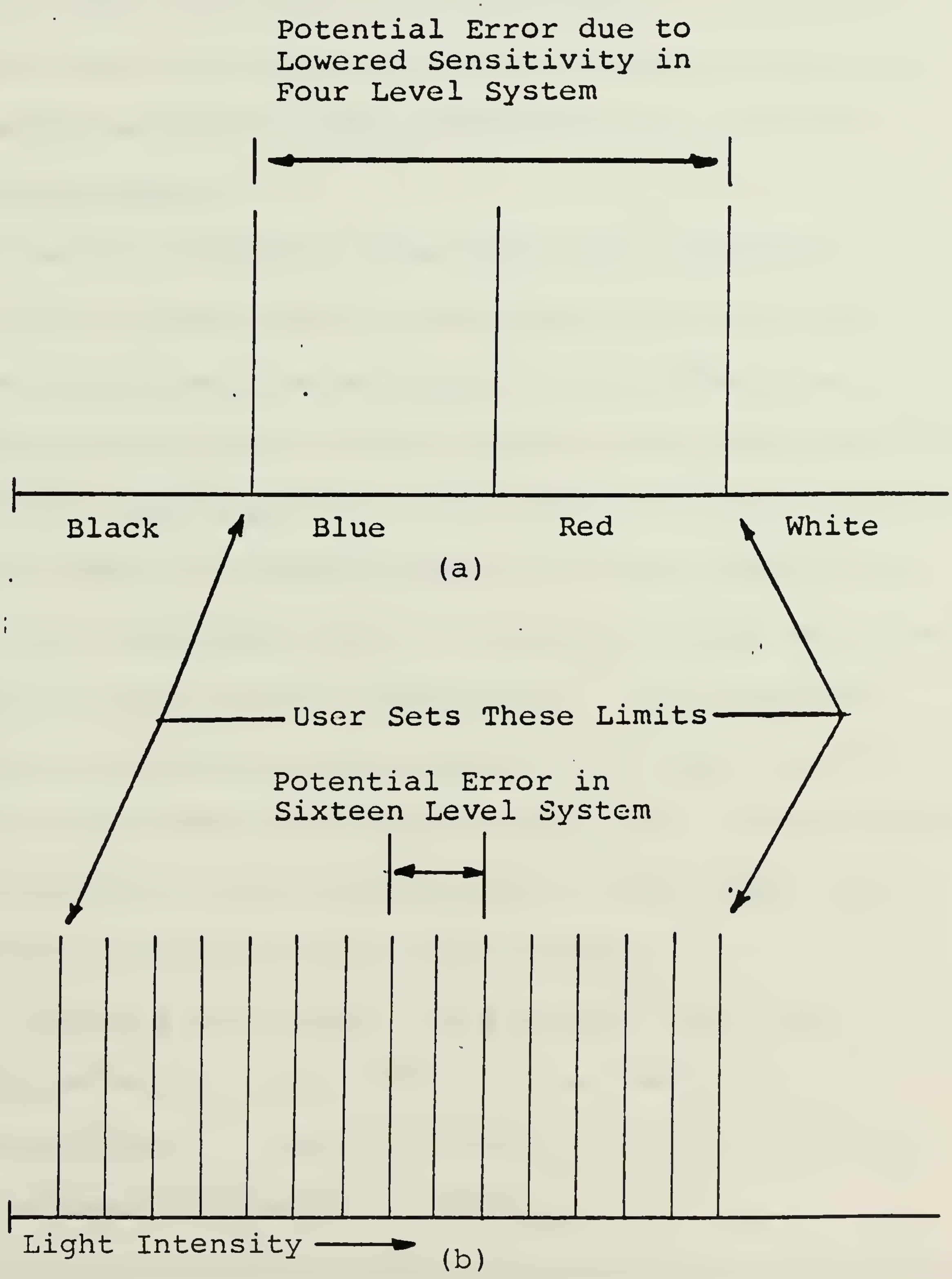
Output

By definition, the hardware has one output: digitized images. The software processes these images, which can be viewed on a monitor, printed on paper, or saved on disk. The value of the system lies in the various pixel counting and image comparison programs.

PCEYE10 is the PCEYE program (in BASIC) whose main utility is to "grab" the analog signal, then display and/or save the resultant digitized images. DIFFPGM, SHOWPIC, and PRNPICS use the images saved by PCEYE10.

DIFFPGM.EXE compares two PCEYE images pixel by pixel. It then reports the number and percentage of pixels that do not match. The sensitivity of the program is set by the user during each run, meaning that the user may choose the range of grey levels that can cancel each other. For example, choosing zero for the sensitivity means only exact matches cancel one another (blue cancels only blue). A sensitivity of one means pixels within one level of each other will cancel (blue could now also cancel black and red). Obviously, zero should be the only feasible setting for this system, since there can be a great difference between the extremes of light intensity of adjacent grey levels in this system. Figure 2.3 illustrates the relative error of a four grey level system to a sixteen grey level system when the sensitivity is set to one.

This assumption of the required sensitivity was tested in the laboratory. From testing, it was found that using zero resulted in a fairly constant mismatch of approximately two percent of the pixels in any two images: 1) grabbed close together in the laboratory; 2) under artificial lights; and 3) directly from the camera. This error is minor and consistent, and therefore of little concern. Then



INTENSITY RANGE OF GREY LEVELS FOR (a) FOUR AND (b) SIXTEEN GREY LEVEL SYSTEMS
FIGURE 2.3

the step of recording these same images onto videotape was introduced. When the images were replayed to the system, the standard error increased to approximately 3.8 percent. Again, though the error nearly doubled, it is relatively minor and consistent.

When images obtained in the field were subjected to the same test, the errors were in the fifteen percent range. This error is no longer acceptable; it is evident that the difference is the result of the uncontrolled characteristics of the natural lighting and is therefore not easily predictable. The change in lighting appears to vary beyond the bounds of one PCEYE grey level. A solution might be to use one instead of zero for the sensitivity. As discussed above, such a decrease in the sensitivity yields results which can be too gross for realistic analysis. This problem will be resolvable upon introductions of more grey levels, where sensitivity can be more finely tuned.

This may be a moot point, as the pixel for pixel matching is effective only where the two images can be absolutely overlaid. The significance of this constraint will be discussed in Chapters 3 and 4.

SHOWPIC.EXE and PRNPICS.EXE exist to display previously saved PCEYE images on the screen, and to print such images, respectively. They serve no direct analytical purpose; however, review of hard-copy images can bring insight to the patterns assigned by the system. This is demonstrated in

Figure 2.4, where portions of the dark brown aluminum panels present a pattern similar to the buff brick.

HISTO takes the complete screen image and produces a histogram of the number of pixels in each grey level. When counting the overall number of pixels per grey level in an image, the need for exact alignment is relaxed, but camera coverage still must be approximately the same. This is in contrast to the exact overlay required when doing the point-by-point comparison described above. In this case, individual pixels may still have the fluctuation noted while using DIFFPGM, but by comparing the whole field of pixels, the variations of individual pixels should tend to balance.

MEDAOI, developed at the Pennsylvania State University, lists the number of pixels in each grey level of an image, or any rectangular portion thereof; the user defines the upper left and lower right corners. The capability of this program to take any portion of the image is a distinct advantage, as areas external to the object can be reduced, but it is extraordinarily slow.

Videotape Equipment & Recording Media

There is a great range of equipment available on the market today for recording video images. The quality of images produced can usually be correlated to the expense of the system; therefore, it will be important to determine the minimum image quality necessary for processing.

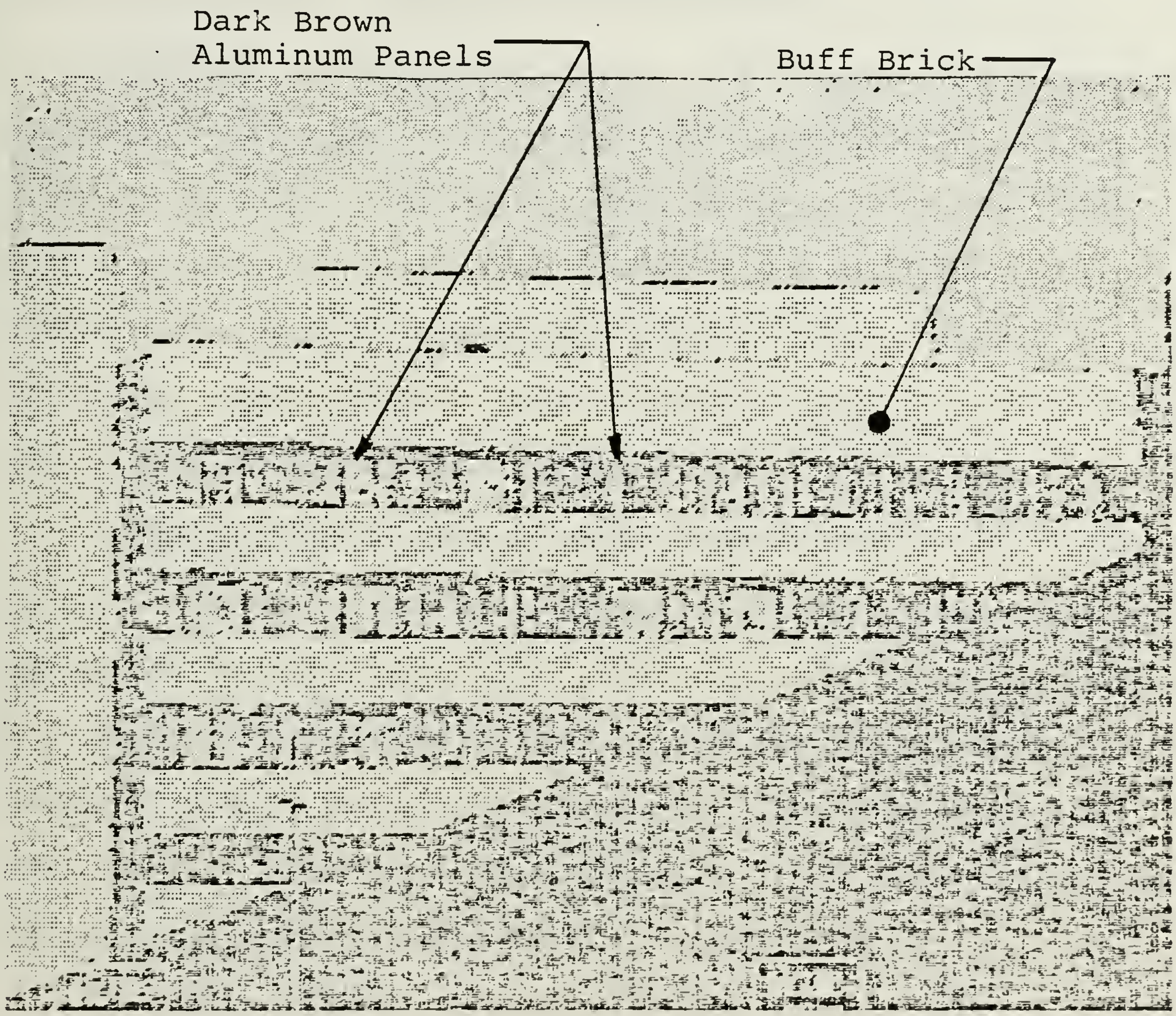


ILLUSTRATION OF SIMILAR DIGITIZED PATTERNS
FOR DIFFERENT MATERIALS
FIGURE 2.4

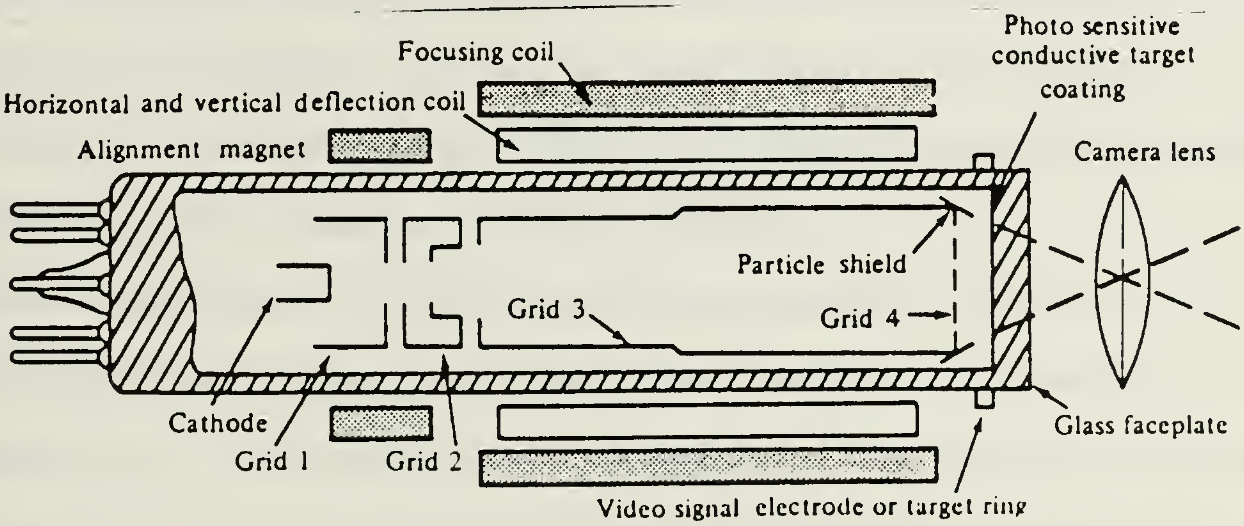
Camera Features: Advantages and Disadvantages

Tube Type (Vidicon) versus Solid State (CCD). There are two general video camera types that are available to the retail purchaser. These are illustrated in Figure 2.5.

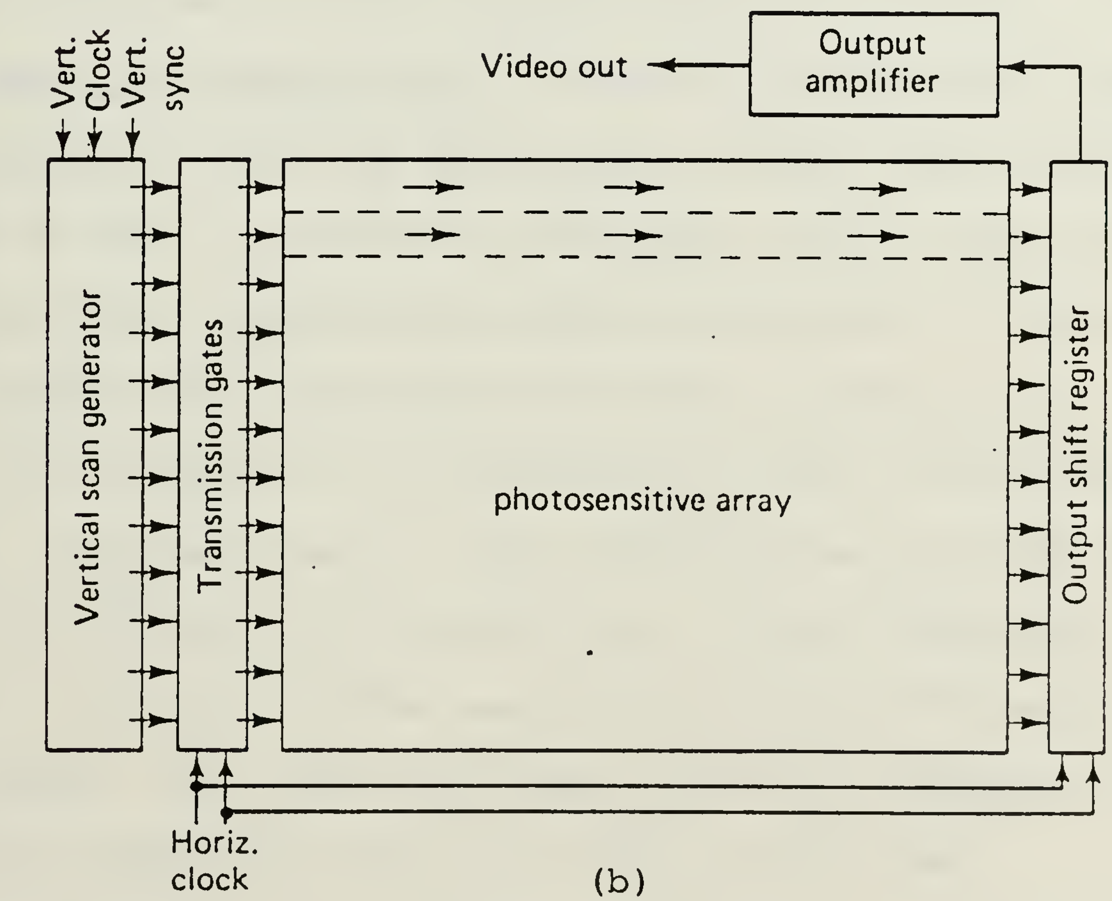
The first is the older, tube type camera, also known as "vidicon." This now conventional device captures the visual image by passing an electron beam across a reactive screen. This produces an analog signal which is combined with a synchronizing signal within the camera, and is output as the RS-170 (or NTSC) signal. This type of camera is currently more prevalent due to its longer time on the market.

Although this is the type of camera used in this research (because it was available), it is not preferable for use on the construction site; the tube used in these cameras is fragile, susceptible to rough handling and environmental extremes, and is subject to "burn-in" (physical damage to the reactive screen) if directed at a light source, such as a window reflecting the sun or a welder at work.

The second, newer type, the solid state camera, is also known as a charge-coupled device or "CCD". This type of camera also has a screen onto which the image is cast, although the screen here is an array of "wells" on the surface of a semiconductor chip. The array is electronically scanned, and produces a digital signal. Unfortunately, this digital signal is then converted to analog within the camera in order to match the current standard (RS-170) for record-



(a)



(b)

(a) VIDICON CAMERA AND (b) CCD CAMERA ELEMENT
(Ballard and Brown, 1982)
FIGURE 2.5

ing on standard videotape and for viewing on a standard television or monitor. Also, as discussed above, this convention has necessitated the use of an analog-to-digital converter in any computer vision system.

The advantages of the solid state camera are its overall hardiness over the tube type, and its increasing availability. The solid state components are not as susceptible as the tube to environmental conditions or burn-in. Manufacturers have almost phased out the tube type camera in favor of the solid state cameras for retail sales.

Zoom lens capabilities. Use of a fixed focal length lens greatly limits the flexibility of any data collection system in that it creates an extremely narrow range envelope from which the required data can be collected. A zoom lens can increase that envelope practically indefinitely.

As an example of the flexibility that a zoom lens provides, the camera used in this research has an eight-to-one zoom (10.5-84 mm), giving the user a large "envelope" within which to set up. In the case of a building 100 feet tall, for example, the usable distance for total coverage of the building ranges from about 150 feet to 1200 feet.

This topic will be discussed further in the treatment of camera placement.

Automatic and manual light level adjustments. Most current video cameras have the capability for automatic light level control. Most of these allow the user to switch

to manual control. There are distinct advantages and disadvantages to each setting.

The automatic light level adjustment feature allows the camera to adjust the iris to meet existing light levels in the environment. This allows the overall lighting levels to change and produce no significant difference in the perceived image as captured by the camera. The major problem with the automatic adjustment occurs when there is a "hot spot" within the field of view, or even a bright object near the camera and not in the picture. The former could be a highly reflective surface, such as a window, or a welder at work; the latter a car in the foreground (as happened during the study). Either causes the camera to close down its aperture because of the perceived light increase, and causes the recorded image to be dark.

The manual adjustment of the light level allows the user to ensure that the light reflected from the surface of interest remains fairly constant. This requires the overall lighting remain constant. If the period of observation is short, and assuming natural light and a relatively clear day, little change in overall lighting is likely. In this ideal situation it is a simple matter to use the fixed, manually set iris found on any standard camera. In the more common situation, it will be necessary to determine independently from the camera what the actual ambient lighting level is, through the use of a properly shielded

and directed light meter, and thereby set the proper iris manually or electronically. The latter system would simply be a customized automatic light level adjustment, but would overcome the problems inherent in the standard through-the-lens system. Again, as discussed previously, minor changes in lighting can be compensated by reducing the sensitivity on a better grey level system.

Videotape Characteristics

Video cassette recorders (VCRs) are designed to produce a color video signal which, while carrying the same information as the standard RS-170 signal, also carries the extra information that is used to produce a color image. As noted above, most frame grabbers are equipped to filter out the extra information.

Resolution. The resolution of a videotaped image appears to be comparable to the direct input of the camera. The experimental results described above, using the comparison program, would seem to support this assumption. The variations in the videotaped image that were imperceptible to the human eye were picked up by the PCEYE system, but were acceptable.

Distortion. Another inherent problem with using the videotape medium is the susceptibility of the tape to wear, causing degradation of the signal output, and the mechanical wear of the VCR mechanism, causing an increase of line

noise. Such degradation and interference can be controlled by limited reuse of videotape, and proper maintenance (especially head cleaning) of equipment.

CHAPTER 3

POSITION OF CAMERA - SITE TESTING

Nature of Observed Object

The nature of the observed object is critical. Its nature is its general shape, dimensions, finishes, edges, depth, and visual accessibility (unobscured).

Humans can observe an object, and subconsciously "fill in the gaps." The computer cannot do this, so it is necessary for the entire image to be captured by the camera or other equipment in use. The ability to capture that entire image is directly related to and restricted by the capabilities of the equipment in use. In this specific case, the equipment is a standard video camera with a zoom lens. The zoom lens allows great flexibility in locating the camera so the entire image can be captured and the effects of distortion caused by being too close to a large object can be minimized. It can also introduce other errors, as will be discussed later in this chapter.

Another factor which will come into play is the nature and orientation of lighting. People have a much more continuous discernment of light levels than the computer, and are therefore more tolerant of poor lighting conditions. Where these poor lighting conditions exist, for instance casting dark shadows, proper analysis will be difficult.

Volume, Area, or Length

A vertical area, such as a wall, provides the greatest ease of observation. The reason for this is the great flexibility available in camera setup positions. As will be demonstrated, the error incurred by varying the position of the camera, within the operational envelope, is minimal and easy to calculate (see Chapter 4 for errors introduced by camera position). This is also true for the error introduced by the distance differences of the extremes of the object. Furthermore, as will be discussed in Chapter 4, it is a simple matter to make a direct correlation between the number of pixels in the digitized image and the actual area observed.

These statements concerning vertical areas can be adapted to horizontal areas, provided that the camera can be positioned within the corresponding operational envelope.

The system may be applied very simply to calculate lengths by establishing the endpoints of the object. Or, if the width of the object is known, the area calculations can be manipulated to yield the width.

The most complex feature to observe is a volume. It is probable that proper analysis of a volume is beyond the capability of a pattern recognition system of the nature now under consideration. Volume calculations are most often associated with earthmoving operations, and are therefore

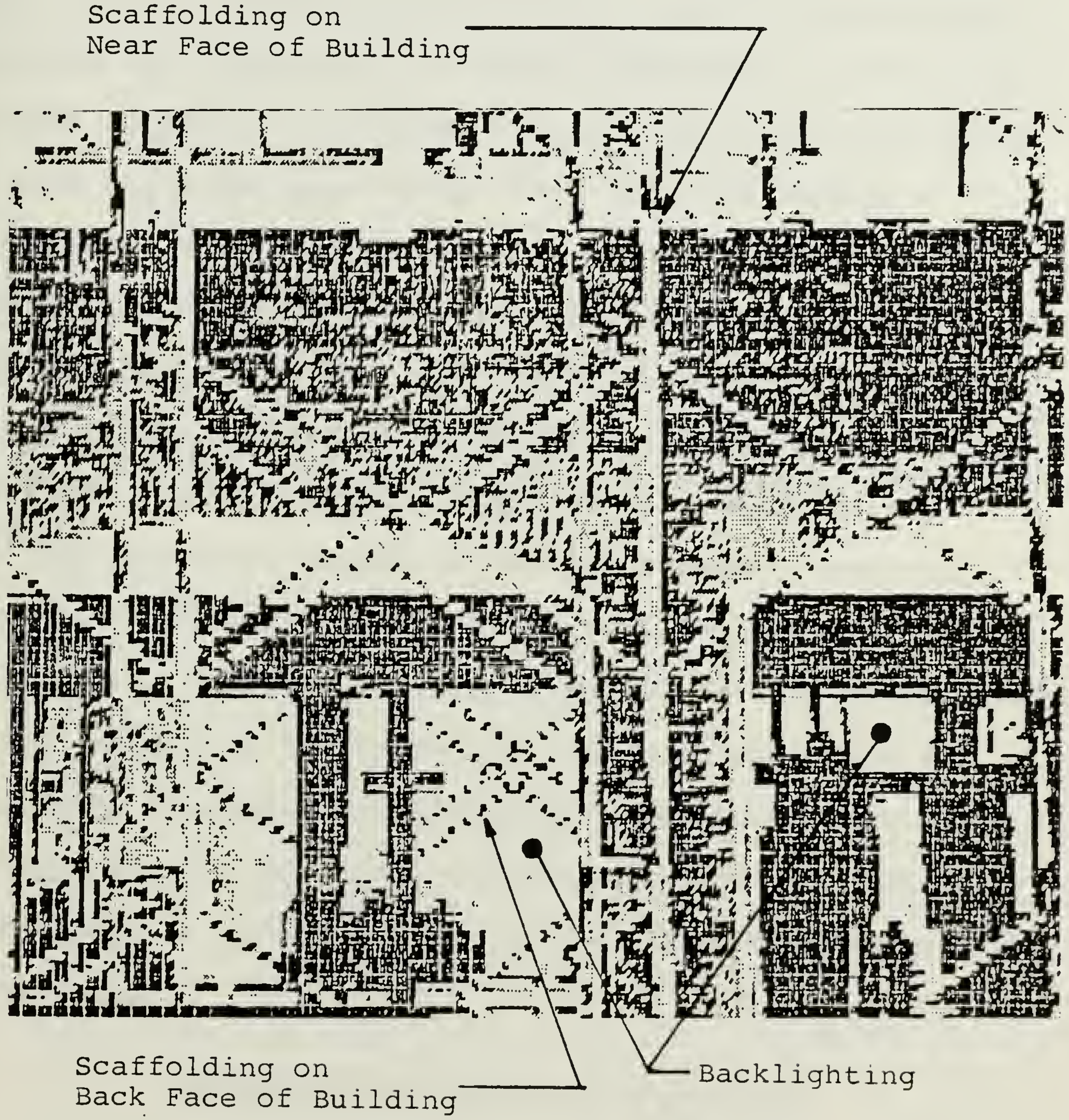
probably in the realm of photogrammetry, at least where the limitation is to visual clues.

Planar or Curved or Noncoplanar

While actual observation of a complex geometry is no more difficult than that of a simple plane, translation of such an image into useful data for measurement of area or productivity, is. The simplest case of a planar object allows direct translation of the digitized image to actual area. When the object is not planar, the user must break the image into segments that the computer will handle individually.

Identification/Description of Selected Structures for Study

Initial observations were collected at construction sites in State College, Pennsylvania. These included a small office building and a hotel. It was determined that these construction sites introduced too many variables at this stage of study. However, the images obtained clearly illustrated the potential and the problems of using a vision system. Figure 3.1 is a PCEYE image showing a portion of a small office building under construction. It was possible to see through to the other side of the building; this is reflected in Figure 3.1 where the scaffolding on the back of the building appears in the image.



BUILDING UNDER CONSTRUCTION SHOWING
EFFECT OF BACKLIGHTING AND SCAFFOLDING
FIGURE 3.1

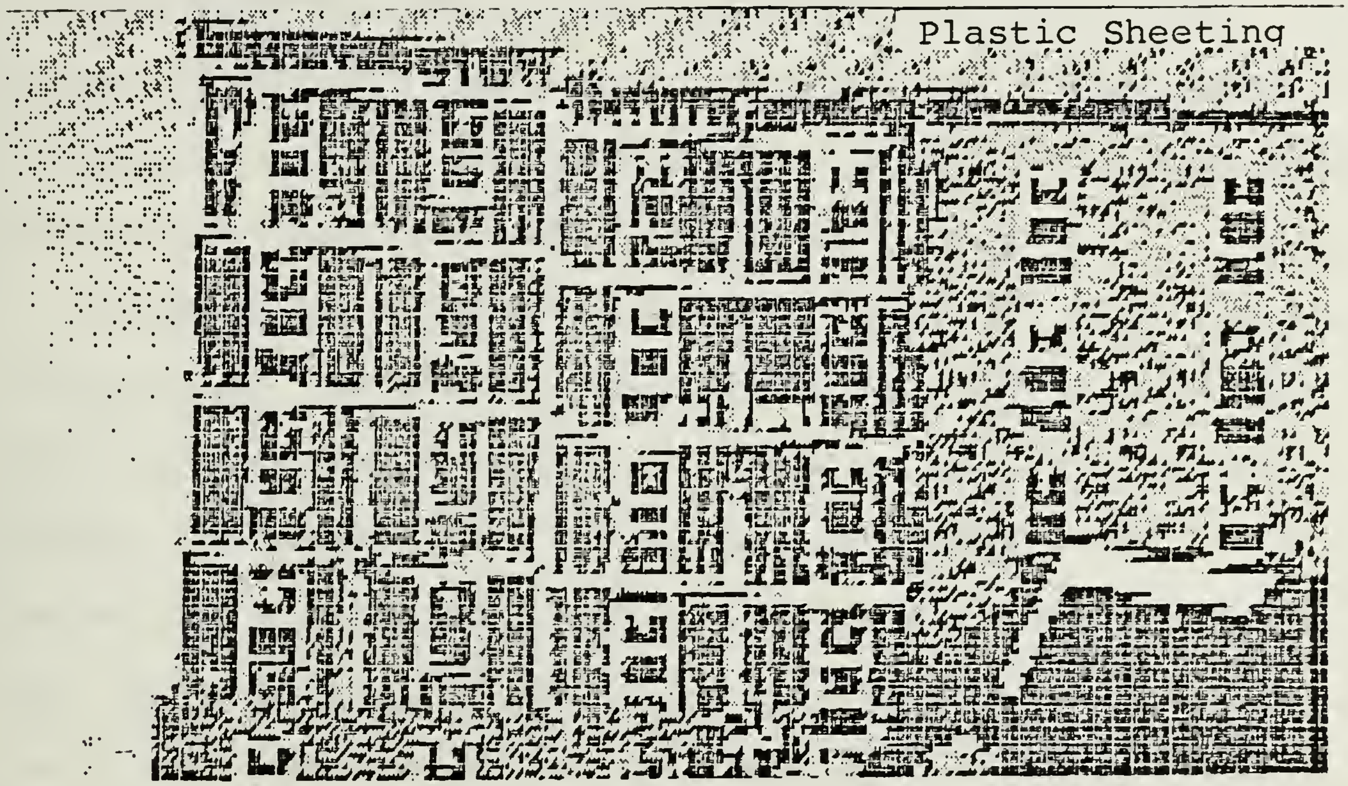
Figure 3.2 shows images developed from the Atherton Hotel, State College, at different stages of its construction. It can be plainly seen that a simple vision system here has the capability to detect differences in the areas of work in place. In Figure 3.2(a), there are large areas of wall that show bare studs, with only the windows installed. The jagged appearance of the roofline in this figure is the result of plastic sheeting, draped over the edge of the roof to keep out the rain, blowing in the wind. Figure 3.2(b) clearly shows that the sheathing has now been installed, and the plastic sheeting has been removed.

The southwest facade of Centre Community Hospital, State College, Pennsylvania, was finally chosen for this study. This subject is of the simplest nature, as described above, in that it is planar, for the most part, and is relatively unobscured by extraneous materials, such as foliage, plastic sheeting, scaffolding, etc. In addition, it provides a limited variety of construction materials for observation.

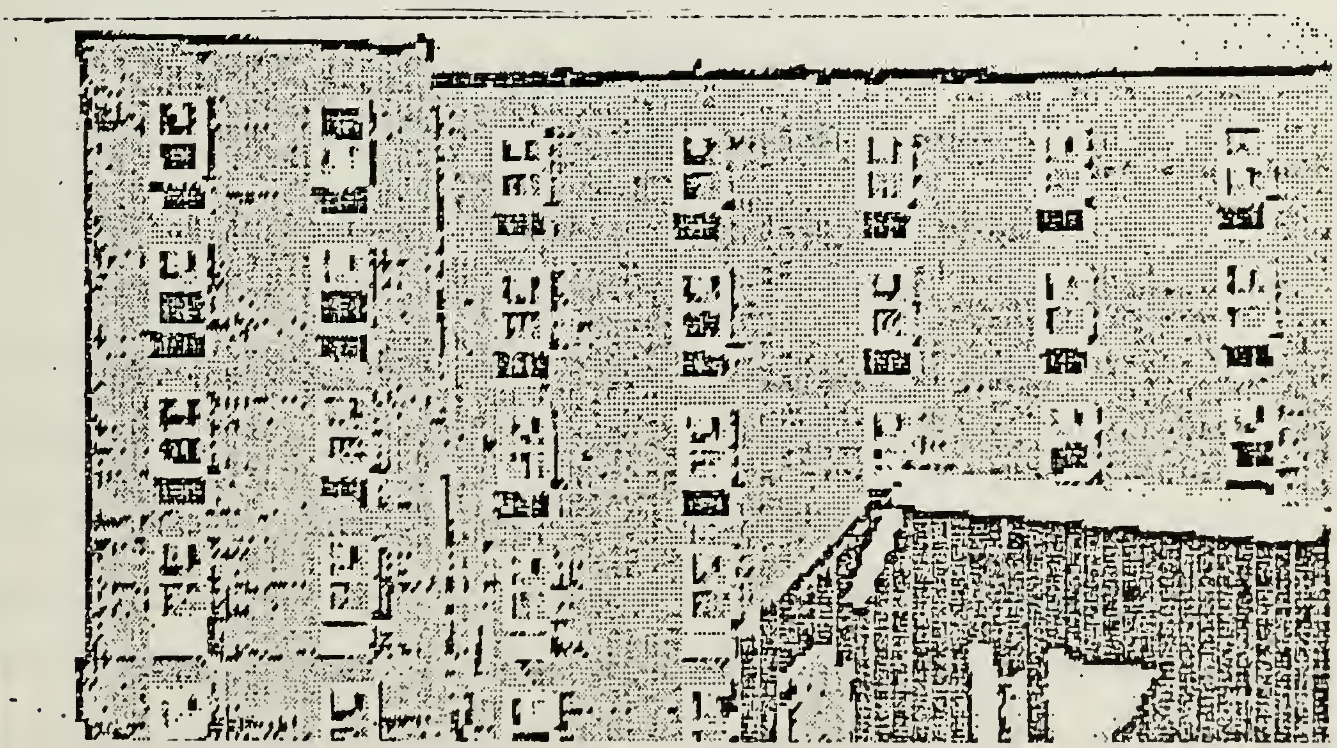
Nature & Condition of Illumination

Natural Light(Sunlight)

Needless to say, the position, intensity, color, and, in fact, existence of natural light are completely uncontrollable. These characteristics can vary greatly from day to day, and <u>will</u> over periods of months.



(a)



(b)

ATHERTON HOTEL UNDER CONSTRUCTION SHOWING
(a) BARE STUD WALLS AND PLASTIC SHEETING
AND (b) CLOSED IN WALLS
FIGURE 3.2

All observations for this project were taken under conditions of natural lighting. As stated, the facade under consideration has a southwest exposure. For reference, the observations were taken in mid-January.

Time of Day. Observations were taken at three times on the same day: 9:00 a.m., 12:15 p.m., and 3:30 p.m. These times were chosen to center the observations between local sunrise and sunset. The 9 a.m. footage shows a distinct shadow line across the facade not evident in either of the other observations. This shadow line moves noticeably in the short (approximately fifteen minute) period of filming. As foreseen above, the computer cannot discern the continuation of the wall into the shadow line; this could introduce a large error over the time span when the shadows move rapidly, i.e. early morning and late afternoon.

Weather. On the day of observation, the weather was officially described as partly cloudy, but was subjectively judged to be a very clear day. Temperatures ranged from the low teens to the mid-thirties (degrees Fahrenheit) through the day. These temperatures were below the recommended operational ranges of the videotape equipment, and although performance was satisfactory, the rechargeable batteries for the system required frequent changing.

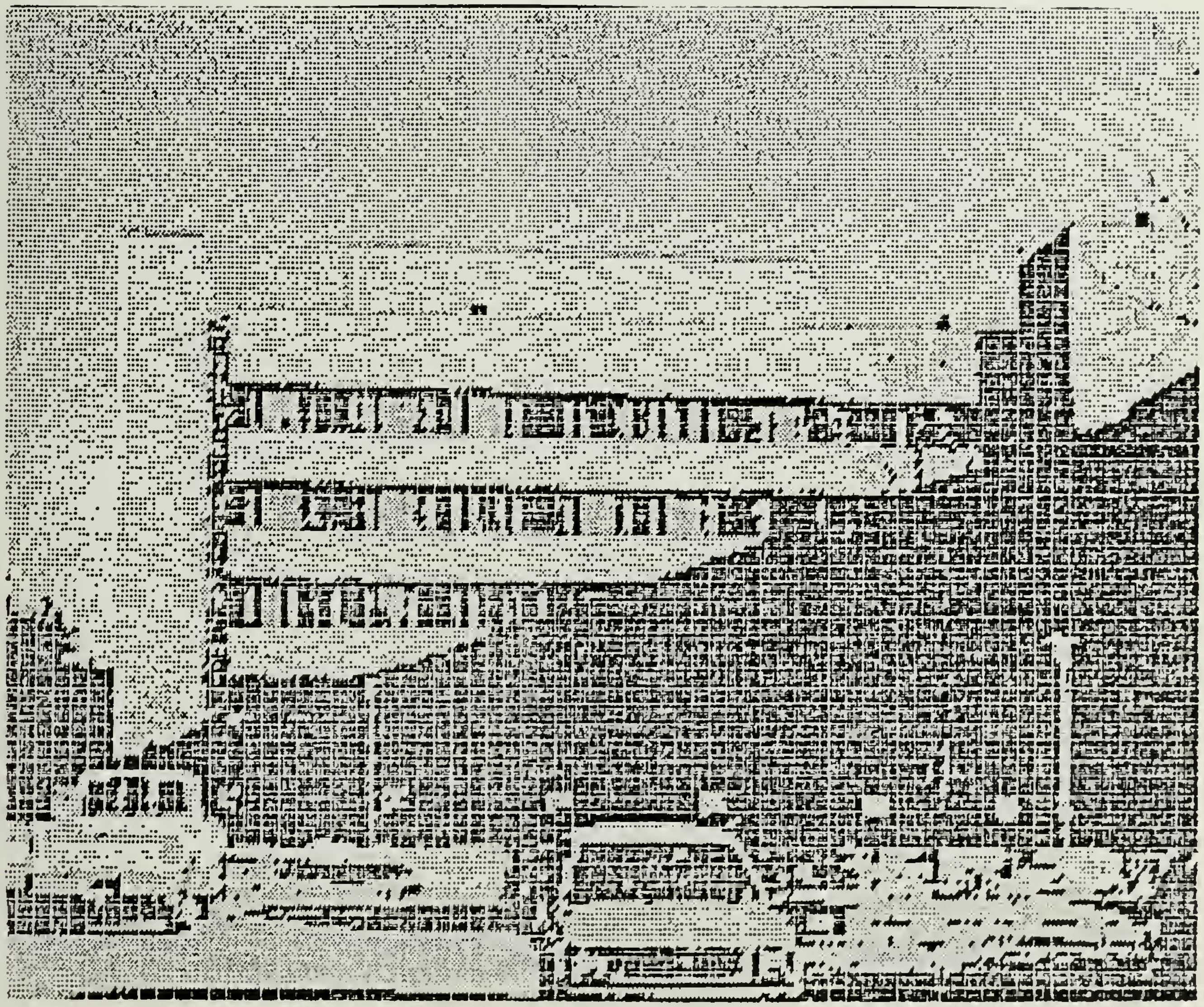
Orientation. The orientation of the sun varies greatly between the three observation periods. The three times of

observation were chosen intentionally to demonstrate the differences introduced by the changed angle of incidence of the sunlight upon the building face. The specifics of how this orientation will change is also dependent upon the time of year.

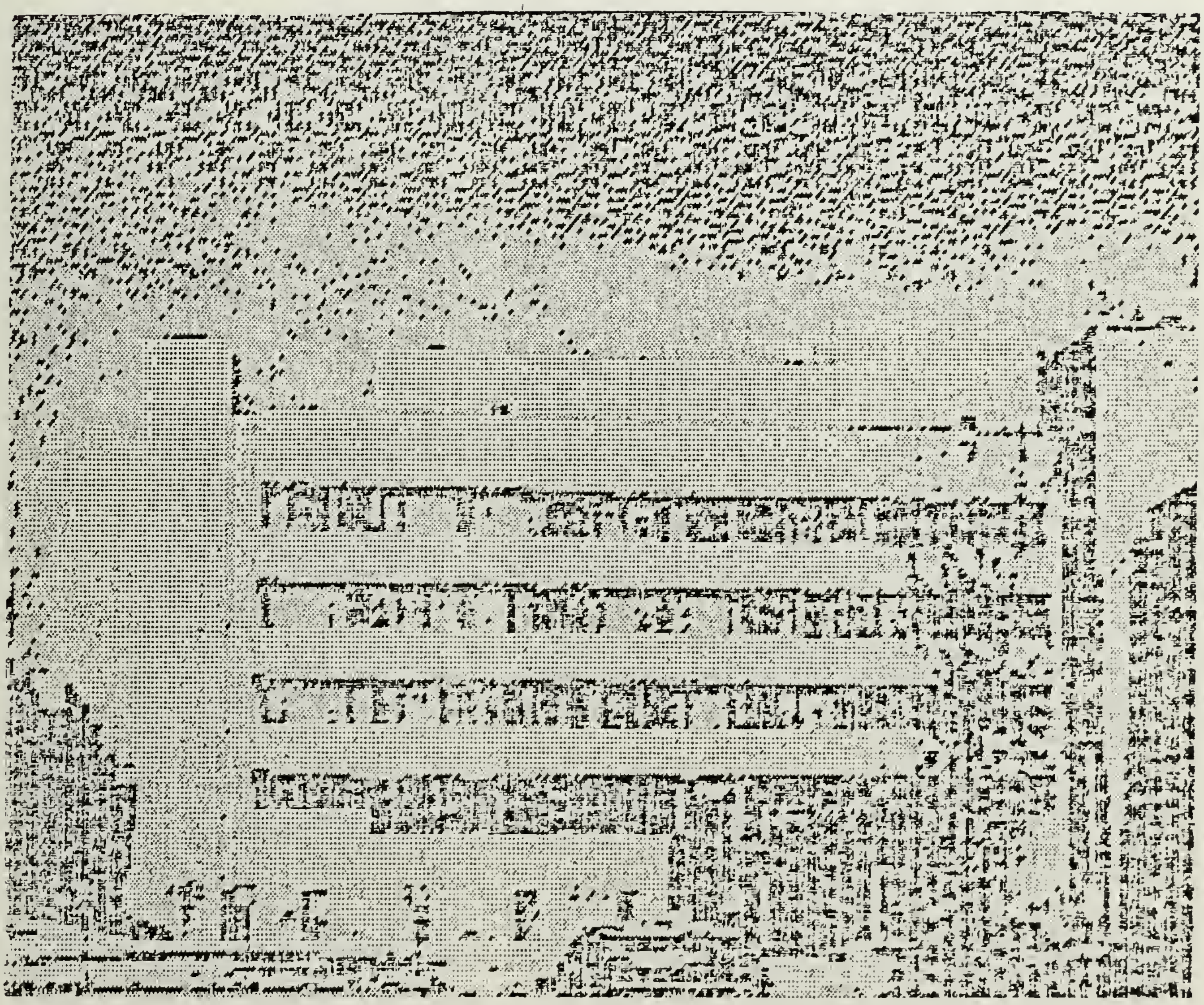
As can be seen in the morning image (Figure 3.3), the sun was low and to the southeast of the building. The adjacent hospital wing, on the right, is casting the large shadow across the observed face. The noon observation (Figure 3.4) placed the sun almost directly behind the observation camera. This caused the reflection from the buff brick surfaces to saturate the image, reducing the contrast. The late afternoon image (Figure 3.5) shows more contrast than the noon image, but the nature of the illumination is also different from that observed in the morning. The overall light level is higher than in the morning image. Additional images from the three observation periods can be found in Appendix A.

Artificial Light.

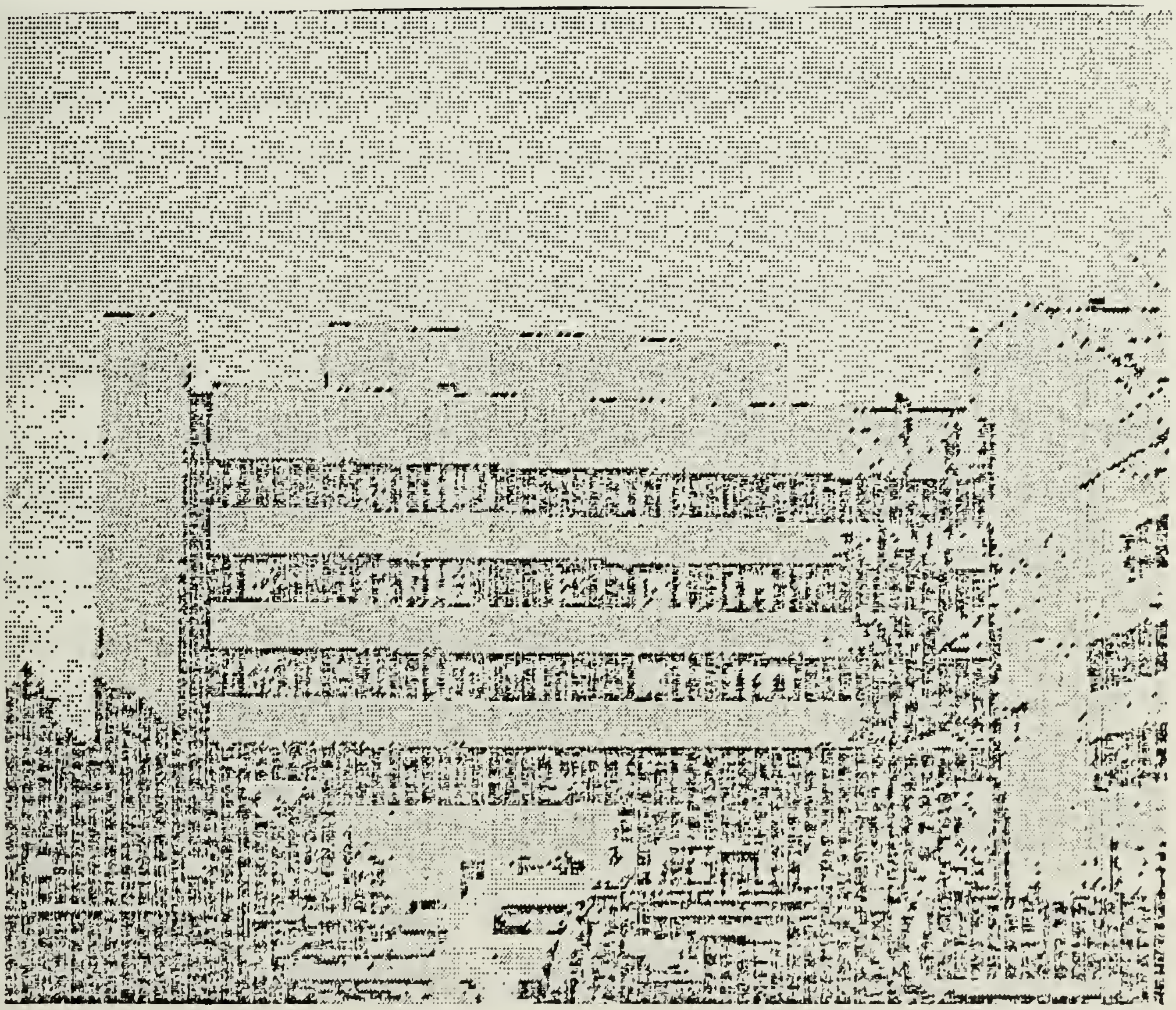
The use of artificial light has both advantages and disadvantages. Its primary advantages are: the light level is constant and controllable; the orientation of the light source is fixed and controllable; and additional lighting can be added to minimize shadowed areas. The primary disadvantages are: artificial lighting is expensive; and it



NINE A.M. IMAGE OF CENTRE COMMUNITY HOSPITAL
SHOWING SHARP SHADOWLINE
FIGURE 3.3



NOON IMAGE OF CENTRE COMMUNITY HOSPITAL
FIGURE 3.4



AFTERNOON IMAGE OF CENTRE COMMUNITY HOSPITAL
FIGURE 3.5

can cause harsh shadow lines which the computer cannot "see" through.

Orientation with respect to object. Control of the orientation of artificial lighting was cited as a primary advantage in the previous section. This, of course, is in contrast to natural light variations and the weather.

There are three overall orientations which can be addressed for the illumination of a building exterior. The first two involve lighting attached to the building itself and directed either up from the bottom, or down from the top. This scheme of lighting is acceptable in a constructed facility where the sole purpose is to show the building. However, in a building under construction, the purpose of illumination is to enable the craftsman to carry out a task. Lighting in the first configuration will likely restrict access to the structure. In the second configuration, the light will likely be in the workers' eyes. In either case, the lighting is not an aide but a potential safety hazard.

The third, and preferable, configuration is for the lighting equipment to be located some safe distance away, and the light directed back onto the structure. The distances used are obviously dependent upon the size of the building, the site conditions, and the capabilities of the equipment. Also, the lighting should be placed at whatever height best illuminates the areas of interest, as is physically possible.

The artificial illumination used to see the exterior of the building is not the only case of concern in capturing images. As can be seen in the prints of the images (Figures 3.3, 3.4, 3.5, and Appendix A), lighting inside the building drastically changed through the course of the day. In the observed case, the changing interior lighting was the result of light-colored curtains in the windows being opened and closed, and reflecting the sunlight. This is an excellent illustration of the corresponding effect that interior lights on a construction site would have on the image. In the worst case, it would adversely effect the overall lighting level, and degrade the usefulness of the image for comparison. Much the same effect can be observed in the situation where unshielded welding is ongoing.

Another form of "artificial" light which should be considered is backlighting. If the building is at a stage where you can see all the way through it, the light coming through is indistinguishable from the light reflected from the structure. This was clearly illustrated in Figure 3.1.

<u>Interference and/or Combination with Available Natural Light</u>. The relationship between natural and artificial light could be most beneficial when the artificial is used to fill in the gaps, or shadows, left by the natural. In bright daylight this is unrealistic, but in the early morning or evening, or in heavily overcast conditions, this could be of great benefit. The gap in this hypothesis is

that these periods already require artificial lighting due to the low natural lighting level.

Camera Focusing Precision

As described earlier, the differentiation program showed a regular error, even when observations were taken very close together, the camera was not moved, and even when the observations were taken under artificial lights in the controlled setting of the laboratory. The error was larger, but still fairly consistent, when the additional variable of natural light in the field was introduced. The basic error level may be the result of inconsistency by PCEYE when assigning the grey scale values to an image. We cannot predict this error ahead of time, but we can determine it under each new set of field conditions, and account for it in our calculations.

The problem in using a differentiation type program is that the compared images must be matched precisely. This known as image registration. Using standard video equipment, daily camera setups are impossible to duplicate. The actual precise placing of a standard video camera is as simple as using a plumb bob; the key for this application is directing the camera so as to precisely duplicate the captured image. For this exact duplication to be achieved, it will be necessary to use methods of precision optical alignment, common in surveying; this must be done through

the lens or through other optics attached to the camera, and referenced to a fixed object. From that reference, it will be necessary to have the capability of rotating the camera in the vertical and horizontal planes to the desired position (as with a transit or theodolite). This will likely entail permanent modification and customization of equipment, as the standard through-the-lens image lacks the precision necessary.

Range Targetable at One Time (Depth of Field)

The primary limitation evident in the depth of field variable of the video camera is the randomness with which the camera focused on the object under wide angle view. For example, at distances up to 300 feet, the automatic focusing mechanism systematically focused at three to six feet, instead of the "infinity" range expected. The visual image does not suffer appreciably, but the actual result is much more homogeneous to the computer. This is acceptable for gross estimates of area, or for simple counting, using the current system. This would not be acceptable for work requiring higher precision, such as applications in quality control.

Proximity to Object Relative to Object Size

The immediate result of collecting data at close range to a large (tall or long) object is the familiar distortion of a tall building narrowing at the top, or the roadway

narrowing in the distance. This distortion decreases as the relative distance from the object increases.

The origin of this distortion is the nature of the "field of view" of the viewing device, this being the human eye or the camera lens. Field of view is expressed in terms of an angle. When this angle is a constant, as it is for any one focal length of a camera, or our eyes, the actual area covered by the field of view is dependent on the distance of the object from the observer. As the distance increases, the actual width or height of the field of view increases proportionally.

In the present case, for illustration, we can assume that the object has constant width. As the distance to the object increases, the object fills less of the field of view. The observer perceives this as the object becoming narrower at the top or in the distance. The greater the difference between the near and far points on the object, the greater the distortion. The geometry of this situation is illustrated and analyzed in Chapter 4.

Role of Zoom Lens at Different Ranges

As briefly stated above, the use of a zoom lens can greatly increase the flexibility with which we can choose a site for the camera setup. It can give us a virtually unlimited number of locations from which the image is obtainable. In addition, it can be a primary tool for

fighting the visual distortion described in the previous section. With a zoom or telephoto lens, the observation can be made from a greater distance, thus removing the distortion, and will still capture the entire image desired.

These advantage of the zoom lens are not without their complications, however. While the flexibility of the zoom lens is ideal for the construction site, where the observer may be required to change position regularly, it also introduces another, imprecisely controllable variable. In the common commercially available video camera, the focal length markings on the lens are vague at best. There are no "click stops" which would indicate precise settings for focal length. The best the user could hope for is that the highest and lowest settings are reasonably reliable; judging by the performance of the model used for data gathering, this is a poor assumption. The zoom mechanism gives a very poor response to manual adjustment.

For these reasons, it is recommended that great effort be made to establish fixed locations from which to gather data, and hence used one of several fixed focal length lenses. If the advantages of the zoom lens are of great need to the user, the zoom must have precise metering for rapid and correct determination of focal length.

CHAPTER 4

PROCESSING OF DATA

Relationship of Pixels to Object Size

The computer does not report the number of square feet. It reports the number of pixels in the digitized image. Users must understand that due to the conditions on the site, it may be necessary to obtain data from different positions. Each camera position will have its own unique relationship to the building, which will determine what the number of pixels equates to relative to the structure being measured. Each unique pixel ratio must be provided to the computer for analysis. There are several alternatives to accomplish this. These alternatives will require the creation or acquisition of software with different capabilities than the system employed in this study.

Manual Input

Manual input is probably the simplest method for informing the computer of the variables applicable to individual sets of acquired images. This requires that specifics relative to the controlling variables be provided at the prompting of the system.

First, some object of fixed dimension within the frame must be identified, from which the computer can derive the ratio of pixels to feet at a given point in the image. This identification can be accomplished by using the cursor keys

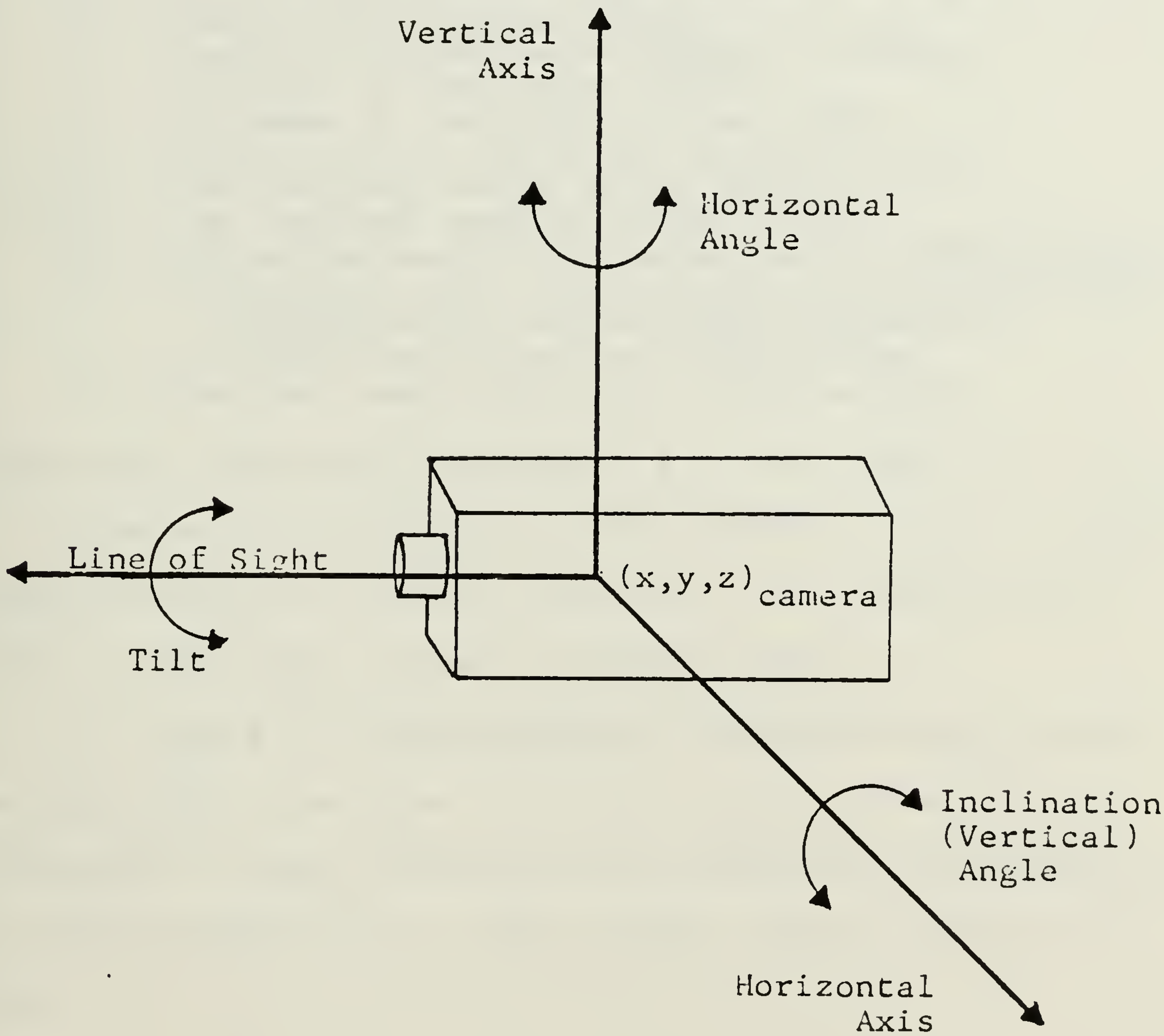
or a "mouse" to establish the endpoints of the object, which could be as simple as a surveyor's range pole. This standard must be at the same distance as the object of the study, or both distances must be known and provided.

In addition, it will be necessary to provide such variables as distance from the object, focal length of the lens, the vertical (elevation) angle of the lens, and the camera relationship to the centerline of object.

The necessity of the first variable (reference object) is straightforward, but qualified in that the ratio determined is in fact accurate only at that point on the screen. Deviation from the determined ratio increases with the distance from that point.

As discussed in Chapter 3, the angular field of view of a camera is fixed, but the actual dimensions covered depend directly upon the distance of the object from the lens. As the distance increases, the field of view grows, so an object of fixed dimension appears to shrink. This explains why a tall building gets narrower at the top.

It is necessary to quantify this effect if a true output is to be obtained from the system. The ratio of pixels to square feet is variable, and dependent upon the location of any particular point on the screen. This is controlled by the six-dimensional relationship of the camera to the object (x, y, z, inclination, horizontal angle, and tilt). This relationship is illustrated in Figure 4.1.



SIX-DIMENSIONS OF CAMERA PLACEMENT
FIGURE 4.1

The first step in defining the equations which will determine the ratio is to define the variables to be discussed:

D_h = horizontal distance from camera to object
h_o = height of reference object
h = height of object with respect to camera
D_o = distance to reference object
D = distance to any point in the image
R_o = screen row of the fixed dimension given
α_v = the vertical angle of the camera
σ_v = the vertical angular field of view
C_o = screen column of the centerline of fixed dimension
α_h = horizontal angle of the camera with respect to perpendicular from building centerline
σ_h = the horizontal angular field of view
δ = angular separation from reference point

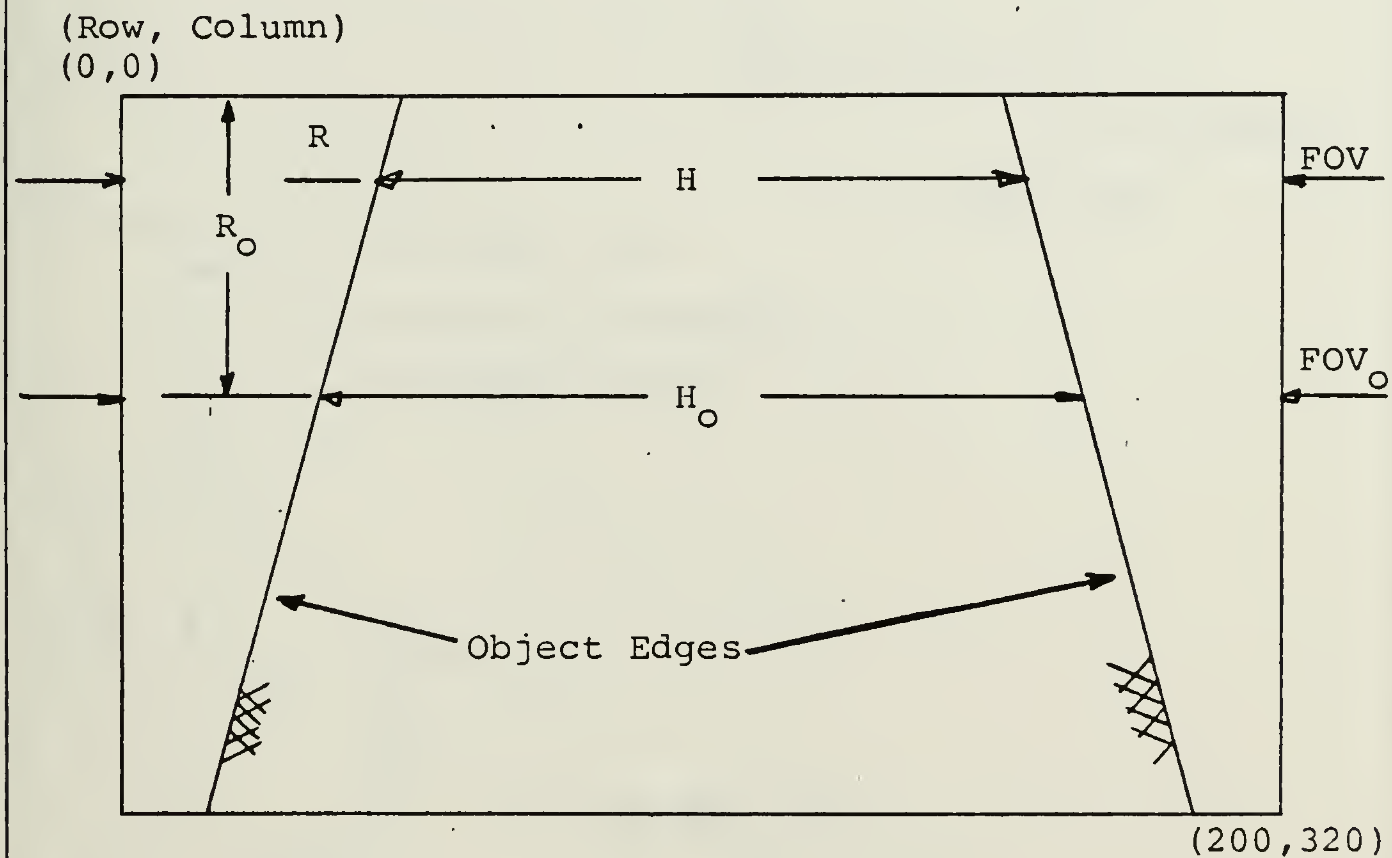
These are illustrated in Figures 4.2, 4.3, and 4.4.

The ratio of the sizes of the fields of view at any two points moving up or down a structure will determine the ratio of the pixel "sizes" at those levels.

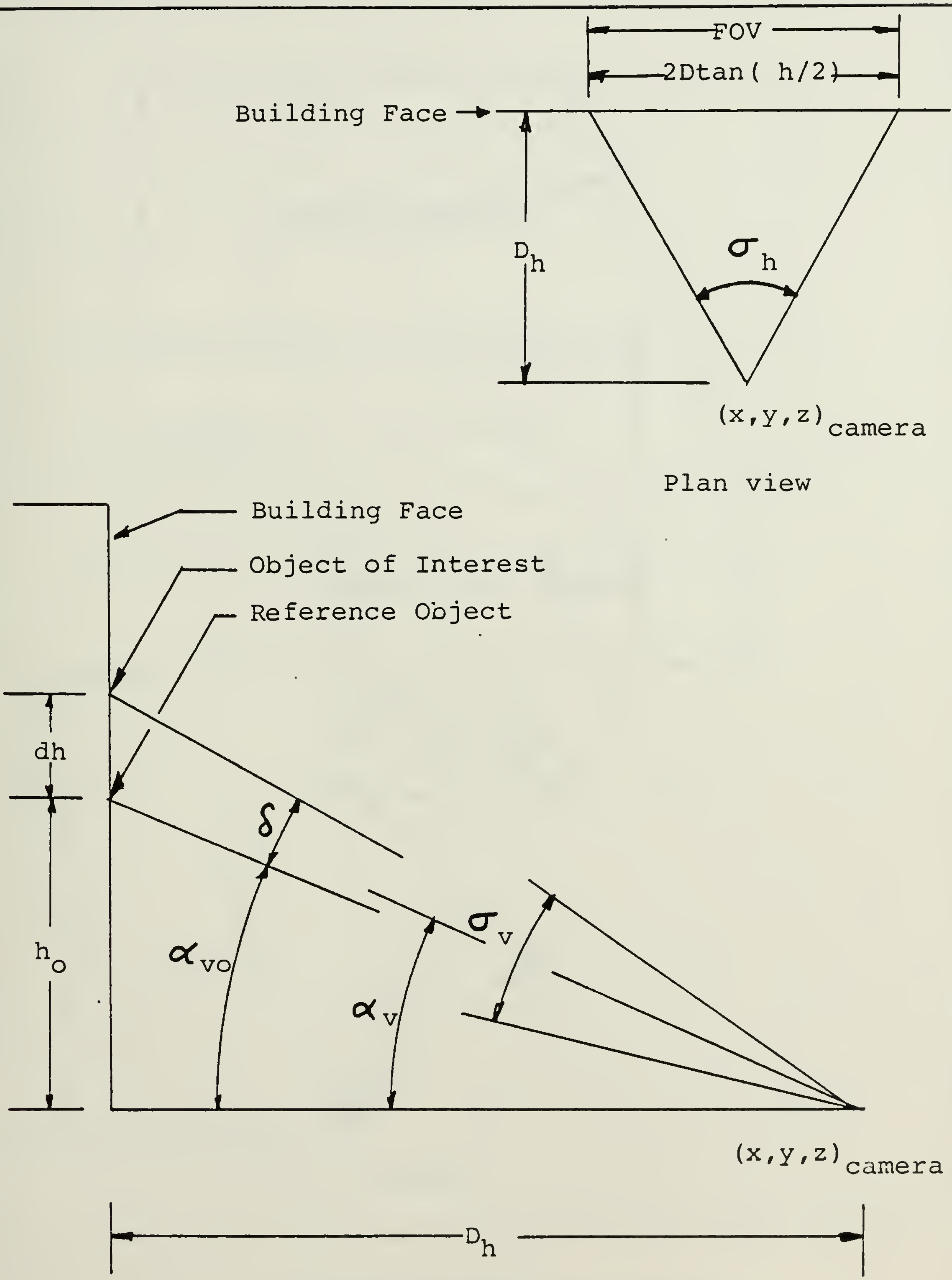
In Figure 4.2, the dimensions designated as H_o and H are physically the same. In this example, however, the distances to those two points are different, designated by D_o and D respectively. These values, shown in Figure 4.3, are:

$$D_o = \sqrt{D_h^2 + h^2}$$
$$D = \sqrt{D_h^2 + (h+dh)^2} \qquad (4\text{-}1)$$

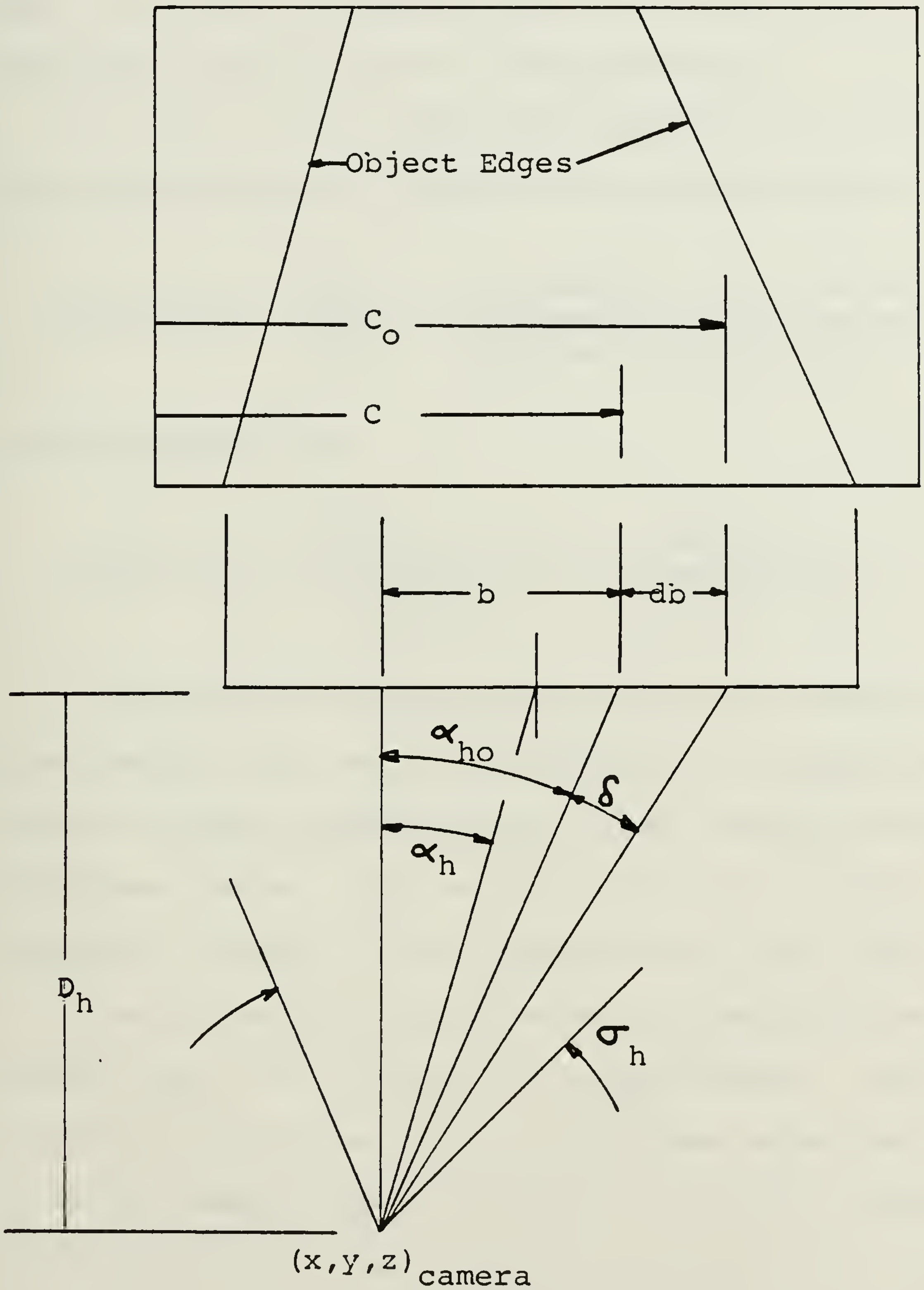
But $h = D_h \tan\alpha_{vo}$ and $dh = D_h(\tan(\alpha_{vo}+\delta) - \tan\alpha_{vo})$; therefore, by substitution,



SCREEN IMAGE ILLUSTRATING NARROWING DISTORTION
AND SCREEN VARIABLES R AND R_o
FIGURE 4.2



PHYSICAL RELATIONSHIPS CAUSING
VERTICAL DISTORTION
FIGURE 4.3



PHYSICAL RELATIONSHIPS CAUSING
HORIZONTAL DISTORTION
FIGURE 4.4

$$D_o = D_h \sqrt{1 + \tan^2\alpha_{vo}} = D_h \sec\alpha_{vo} \qquad (4\text{-}2)$$

and $$D = D_h \; 1 + \tan^2(\alpha_{vo} + \delta) = D_h \sec(\alpha_{vo} + \delta).$$

Since the field of view is represented by

$$FOV = 2\ D \tan(\sigma_h/2) \qquad (4\text{-}3)$$

the ratio of the two fields of view shown above is

$$\text{FOV ratio} = \frac{FOV}{FOV_o} = \frac{2\ (D_h \sec(\alpha_{vo} + \delta))\ \tan(\sigma_h/2)}{2\ (D_h \sec\alpha_{vo})\ \tan(\sigma_h/2)} \qquad (4\text{-}4)$$

which simplifies to:

$$\text{FOV ratio} = \frac{\sec(\alpha_{vo} + \delta)}{\sec(\alpha_{vo})} = \frac{\cos(\alpha_{vo})}{\cos(\alpha_{vo} + \delta)} \qquad (4\text{-}5)$$

This field of view ratio, which will also be termed the Pixel Ratio, will act as a multiplier. By equating the vertical screen dimension to a visual angle, the angular difference between the various points of interest, δ, can be expressed in terms of the screen row of the latter (D).

The system specifications will determine whether the camera image will fill or fit on the screen. Vertical and horizontal directions must be addressed separately. PCEYE clips the image in both directions. For the purpose of this derivation, σ_v has been chosen as the vertical angle of equivalence, as if the vertical dimension of the image has been made to fill the screen vertically.

$$\delta = \frac{R_o - R}{200} \sigma_v \qquad (4\text{-}6)$$

The field of view ratio now becomes the Pixel Ratio by substitution for δ in the FOV ratio equation (4-5).

$$PR_v = \text{PIXEL RATIO} = \frac{\cos(\alpha_{vo})}{\cos(\alpha_{vo} + \dfrac{R_o - R}{200} \sigma_v)} \qquad (4\text{-}7)$$

The v subscript indicates that this pixel ratio adjusts for the distortion due to the differing distance in the vertical direction. A similar derivation, for the horizontal direction, is contained in Appendix B. The variables involved are illustrated in Figure 4.4. C_o represents the centerline of the reference known dimension. The resultant horizontal pixel ratio is:

$$PR_h = \frac{\cos(\alpha_{ho})}{\cos(\alpha_{ho} + \dfrac{C - C_o}{320} \sigma_h)} \qquad (4\text{-}8)$$

The total pixel ratio, the product of PR_v and PR_h, then gives the comprehensive multiplier that can be applied to any point on the screen and determine the physical area to which it equates, based on its (Row,Column) coordinates. In order to put PR_T in terms of the given parameters, it must be recognized that:

$$\alpha_{vo} = \alpha_v + \frac{\sigma_v}{2} - R_o\left(\frac{\sigma_v}{200}\right) = \alpha_v + \frac{100 - R_o}{200}\sigma_v \qquad (4\text{-}9)$$

and $$\alpha_{ho} = \alpha_h - \frac{\sigma_h}{2} + C_o\left(\frac{\sigma_h}{320}\right) = \alpha_h + \frac{C_o - 160}{320}\sigma_h \qquad (4\text{-}10)$$

Substituting into the PR_v and PR_h equations, PR_T becomes a function of Row and Column:

$$PR_T = \frac{\cos\left(\alpha_v + \dfrac{100 - R_o}{200}\sigma_v\right) \cos\left(\alpha_h + \dfrac{C_o - 160}{320}\sigma_h\right)}{\cos\left(\alpha_v + \dfrac{100 - R}{200}\sigma_v\right) \cos\left(\alpha_h + \dfrac{C - 160}{320}\sigma_h\right)} \qquad (4\text{-}11)$$

In this section, it has been demonstrated that the computer may calculate areas by counting the necessary pixels and applying their unique multipliers to obtain a measure of the actual area. That multiplier can vary widely with the conditions and the values of the multiple variables. Appendix C contains a short computer program and several runs showing the range of multipliers for different scenarios using the equipment available. Searching a range of zero to forty-five degrees for α_v and α_h, the maximum PR_T was 2.6818173 and the minimum 0.23047097. It should be noted that this is the worst case scenario, and resulted at both values equal to forty-five degrees, and the lens at its minimum focal length (maximum σ_h, σ_v).

"Automatic" Input

The previous section discussed the steps to be taken by the computer, given a required input by the user. It would be preferable if the user did not have to repeatedly take manual measurements, and then report them to the computer. There are several methods that may preclude this necessity by automating part or all of this particular step in the data collection process. These include incorporation of: radio frequency (RF) triangulation for camera position; a fixed three dimensional target; and utilization of "total station" surveying technology.

RF Triangulation. Simple methods of triangulation can be used to pinpoint the location (x, y, z) of the camera. This data could be stored directly on the videotape and read by the computer at the beginning of any session.

Three Dimensional Target. A three dimensional target of known size, location, and features would serve the purpose of providing a sufficiently intelligent and discriminating system with all the necessary variables described above except for one. It would still be necessary for the range or the focal length of the camera to be provided separately. This separate data input may also be automated through electronic reporting of focal length or electronic rangefinding.

"Total Station" Technology. A total station is an electronic surveying instrument that has the capability to

electronically record all measurements taken. The (x,y,z) coordinates of the station can be determined by initiating any equipment setup with readings taken of points of known coordinates and elevation. The horizontal and vertical angles used to sight these points act as a base of reference for all future observations. The distance to the object is determined by the total station using infrared rangefinding.

It is feasible for our camera equipment to be mounted in concert with the total station so that the two move as one, the camera being a known distance above the total station reference.

How Does Productivity Come From This?

At this point it is necessary to review and address the origin of this specific research project. The purpose intended for the introduction of computer vision to construction was for automated collection and processing of productivity data.

We now know that the information necessary can be collected on videotape and fed to a computer that should be able to determine the actual dimensions of the object being observed. This necessarily includes the capability to have the video input directly to a computer on sight. This has the advantage of providing more immediate data for use.

Determination of straight units-per-hour productivity is a simple matter of the computer making observations at

determinable intervals and comparing, or subtracting, the characteristics of subsequent images. Again, this can be done real-time on the job site, or back in the office at normal speed or fast-forward.

It is possible that with the proper feedback and software, the computer will be able to carry out the operation with little or no supervision. The overriding constraints to this possibility at this time is the limitations on the computer vision system to properly discern between different areas that may provide similar digital patterns. This was observed in several instances among the images obtained for this project.

For the foreseeable future, until the systems available for use are more refined, and the software is developed, these systems will require constant involvement and input by human operators.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Summary

Initiative

The driving motivation behind the entrance into this field was the extremely labor intensive methods required for productivity data collection. The current methods involve manual collection of data which require extensive review and correlation before meaningful results can be determined.

A system is desired with the potential to provide real-time productivity analysis, whereby the supervision on the job site can more rapidly respond to productivity problems, and eliminate those periods of low productivity which cannot be addressed due to the gap between performance and analysis of data.

The potential for automation of this process can be seen in light of the advances in the field of computer science and, more specifically, pattern recognition. This potential has accomplished much in the manufacturing field related to high speed quality control inspection of products on the assembly line.

Approach

The basic approach taken in this research project was to highlight those features unique to the construction site, and to explore the potential barriers in physical characteristics and technology. The predominant factors under consideration are the matters of lighting, distance, and visibility of object.

The initial step was to become familiar with the concepts of computer vision and the basic equipment available for use.

Observations were then taken at several construction sites. However, it was found that the dynamic nature of the facilities under construction introduced too many variables for proper analysis at this time.

After these initial construction site attempts, an existing structure was settled upon for the remainder of the study. This existing facility provided a limited number of materials and a fairly unobscured view, thus reducing the variables in the equation. Observations were then made from three separate distances, at three times of day, at two focal lengths, to determine and illustrate the effects of the differences in lighting (intensity and orientation) and distance (under varied magnification).

Major Problems Encountered

There were many problems encountered during the process of applying the system to the building environment. Tables 5.1 and 5.2 contain a summary of the physical and technological barriers encountered, respectively. The information was put in this format in order to provide the reader a succinct and complete overview of the perceived problems. These Tables should be able to stand alone.

TABLE 5.1

Physical Barriers Encountered

Obstructions	The existence of physical obstructions throughout the printed images is readily apparent. These include trees and shrubs, signs, and lampposts.
Suggestions	The least disruptive method for removing these obstacles from the image is to move inside the obstacles; this is not always possible.
	Physically remove the obstacle. In most cases this may be infeasible, in the case of large equipment, or unpopular, as for trees.
	Develop image "subtraction" routine which can recognize an obstacle as part of the "landscape" and can filter out the obstacle and fill in the remaining gaps by "painting" in the missing object.
Shadows	Shadows drastically change the characteristics of the object being analyzed. Deep shadows caused by bright lighting conditions are opaque/black to the digitizing system.
Suggestions	Utilize vision system with greater number of grey levels. These systems more discriminately assign grey levels and may have the depth to assign non-black values to the shadows.
	Fill in the shadows by supplementing natural lighting with artificial lighting. This will alleviate the severe underlighting that causes the sharp shadow and makes the image black in that region.
Incident Lighting	The angle of the incident light, directly dependent upon the time of day and year, or the location of the artificial lighting, may cause highly reflective surfaces to appear much brighter than they actually are, or to appear to be other materials altogether.

(continued on next page)

TABLE 5.1(continued)

Suggestions	Incorporate color capabilities into the system to better distinguish individual fields of interest.
	Introduce polarization to reduce the intensity of light reflecting from polished surfaces. This will have lesser effect on rough surfaces where the light is reflected more randomly.
	Study surface treatment characteristics of construction materials.
<u>Changing Features</u>	Items such as welding shields, scaffolding, or other temporary facilities which are moved during construction will randomly introduce error into the image.
Suggestions	Manually compensate for the error.
	Develop algorithms which can recognize such obstructions and essentially subtract them from the image, filling in the gaps left behind with the building.

TABLE 5.2
Technological Problems

Optical Distortion	This distortion can be of two types: the narrowing caused by the expansion of the field of view with distance (e.g. on a tall building); and the curving found at the edges of an image caused by imprecision on the part of the camera lens.
Suggestions	The edge distortion is a direct function of the construction and quality of the camera lens; therefore this error can be eliminated by use of higher quality equipment. The narrowing effect of distance can be corrected by use of special lens additions that straighten the inclined lines. The narrowing effect can be compensated through the use of algorithms which define the relative values of pixels at different locations on the screen.
Insensitivity Of Four Grey Level System	The grossness of a four grey level system results in images which can be ill-defined and can exhibit some randomness in assignment of grey levels to regions with comparable light intensities.
Suggestions	Move up to a system with more grey levels. The more grey levels the system has, the narrower will be the actual range of light intensities that will be assigned to any one level. This will allow flexibility in sensitivity definition on a case by case basis. Sixteen or sixty-four grey level systems should be a sufficient next step.
Precise Camera Position and Direction	When the camera position and alignment cannot be permanent or preserved from day to day, the existing equipment does not have the capability for precision in spatial positioning (x, y, and z) or in alignment (elevation, direction, and tilt). The through-the-lens optics are inadequate for these purposes.

(continued on next page)

TABLE 5.2(continued)

Suggestions	Customize camera equipment by mounting on a calibrated base on a fully adjustable tripod, and attaching additional precision sights for referencing established points.
	Mount camera in conjunction with existing "Total Station" technology to determine actual position, rather than attempt to establish a desired position, and translate the new images to the old position algorithmically.

Conclusions

PCEYE as a Learning Tool

The PCEYE system used in the research served as an excellent tool to learn the basics of how a computer vision system works. It also helped to spark an interest in what the potential for such a system might be through showing where its shortcomings were. These limitations are, however, what keeps it from being a feasible system for end use in this effort.

The primary limitations of the system are its low resolution, small number of grey levels, and incredibly slow processing speeds. The first two deficiencies are simple to overcome, as systems exist with higher resolution and more grey levels; they are also quite a bit more expensive. The latter is as much a function of the hardware (a standard PC) as the PCEYE software. Again, this can be overcome by upgrading to currently available equipment, at much greater cost.

Adequacy of Video Resolution

The resolution of the video images obtained were adequate for the purpose of this research, but for actual practice it may be desirable to use more precise equipment, particularly in the areas of focusing and focal length determination in a zoom application. It must be kept in mind that the video camera and equipment used were marketed

for home movies for human observers, and are not expected to produce as precise an image as we desire for our automated processing methods.

Feasibility

The successful construction of a system that automates the collection and processing of visual data for the purpose of productivity analysis is the goal. This will require the extended attention and cooperation of experts in the fields of computer science (and expert systems), pattern recognition, vision systems, and construction and productivity.

The long range goals of this research are feasible. This should be qualified, however, by saying that it will not be done overnight.

Recommendations for Future Direction

There are several areas that warrant immediate consideration as the next steps into adapting vision systems to the construction industry are considered. The overriding and required preliminary step to any of these items, however, is the specific requirement for a new "level" of equipment to be acquired. This equipment should have higher resolution, higher discrimination of light intensities, and higher speed.

The first of these "next steps" or items of consideration is the effect or contribution that the different colors

and finishes of construction materials make to the production of an image. As shown in Chapter 3 and Appendix A, similar patterns in the processed image can be produced by different materials. By studying the effects of colors and finishes, in conjunction with incident light characteristics, the error involved in equating different materials can be recognized and controlled.

The second item which bears investigation is the problem encountered in trying to see into the shadows. It is hoped that a more discerning grey scale breakdown will alleviate this problem.

The third item that was not covered with any depth to this point is the effects of physical obstructions on the construction site, particularly those, such as scaffolding, plastic sheeting, and other temporary structures whose position is constantly or periodically changing. Immovable objects, such as trees, and other buildings, may be compensated throughout the study, but when the obstruction is mobile, the compensation can no longer be manual due to the time involved. The comprehensive effect on the image has not been determined at this time.

The fourth, and probably biggest, step for consideration is the introduction of this system to interior areas of the construction site. To date this work has been directed at inspecting exterior walls. It is another matter altogether for a system to inspect an interior setting where

distances are much smaller in proportion to the items to be observed, requiring a wider angle lens (extremely short focal length). In addition, the camera cannot be stationary and still obtain the necessary images, but its positions must still be determinable for subsequent images to be compared.

BIBLIOGRAPHY

Ballard, D. H. and Brown, C. M., Computer Vision, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1982.

Carlson, A. B. and Gisser, D. G., Electrical Engineering Concepts and Applications, Addison-Wesley Publishing Company, Reading, Massachusetts, 1981.

Dillman, R. J., "The Development of a Methodology for Time and Motion Studies of Scraper Operations", Master of Engineering Paper in Civil Engineering, The Pennsylvania State University, July 1981.

Fowles, G. R., Introduction to Modern Optics, Holt, Rinehart and Winston, New York, 1975.

Kolman, B., Elementary Linear Algebra, Macmillan Publishing Company, New York, 1977.

Paulson, B. C., "Automation and Robotics for Construction," Journal of Construction Engineering and Management, Vol. 111, No. 3, September 1985.

Thomas, H. R. and Smith, G. R., "Conceptual Design of an Automated Real-Time Data Collection System for Labor-Intensive Construction Activities," Proceedings, Fourth International Symposium on Robotics and Artificial Intelligence in Building Construction, Haifa, Israel, June 22-25, 1987, pp. 849-862.

Ward, C. J., "Automatic and Adaptive Controls for Construction Equipment," paper 750765, SAE Off-Highway Vehicle Meeting, Milwaukee, Wisconsin, September 8-11, 1975.

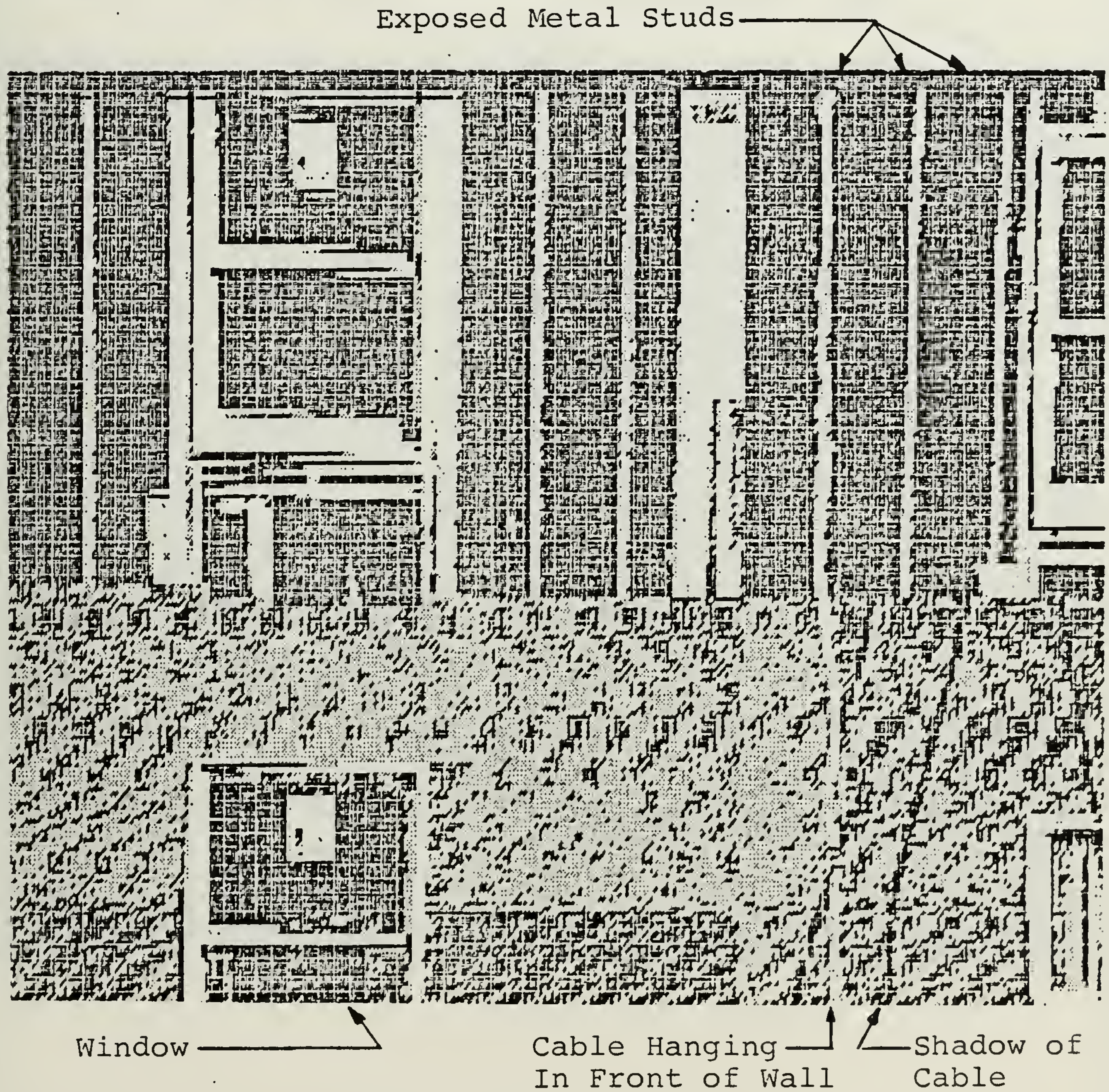
Ward, C. J., "Computer-Aided Planning, Design Estimating and Scheduling of Earthwork," SAE technical paper 831307, International Off-Highway Meeting & Exposition, Milwaukee, Wisconsin, September 12-15,1983.

Ward, C. J. and Springston, P. S., "Field Trials of Laser Surveying and of Experimental Earthmover Blade Control Kit," paper 770551, SAE Earthmoving Industry Conference, Central Illinois Section, Peoria, Illinois, April 18-20, 1977.

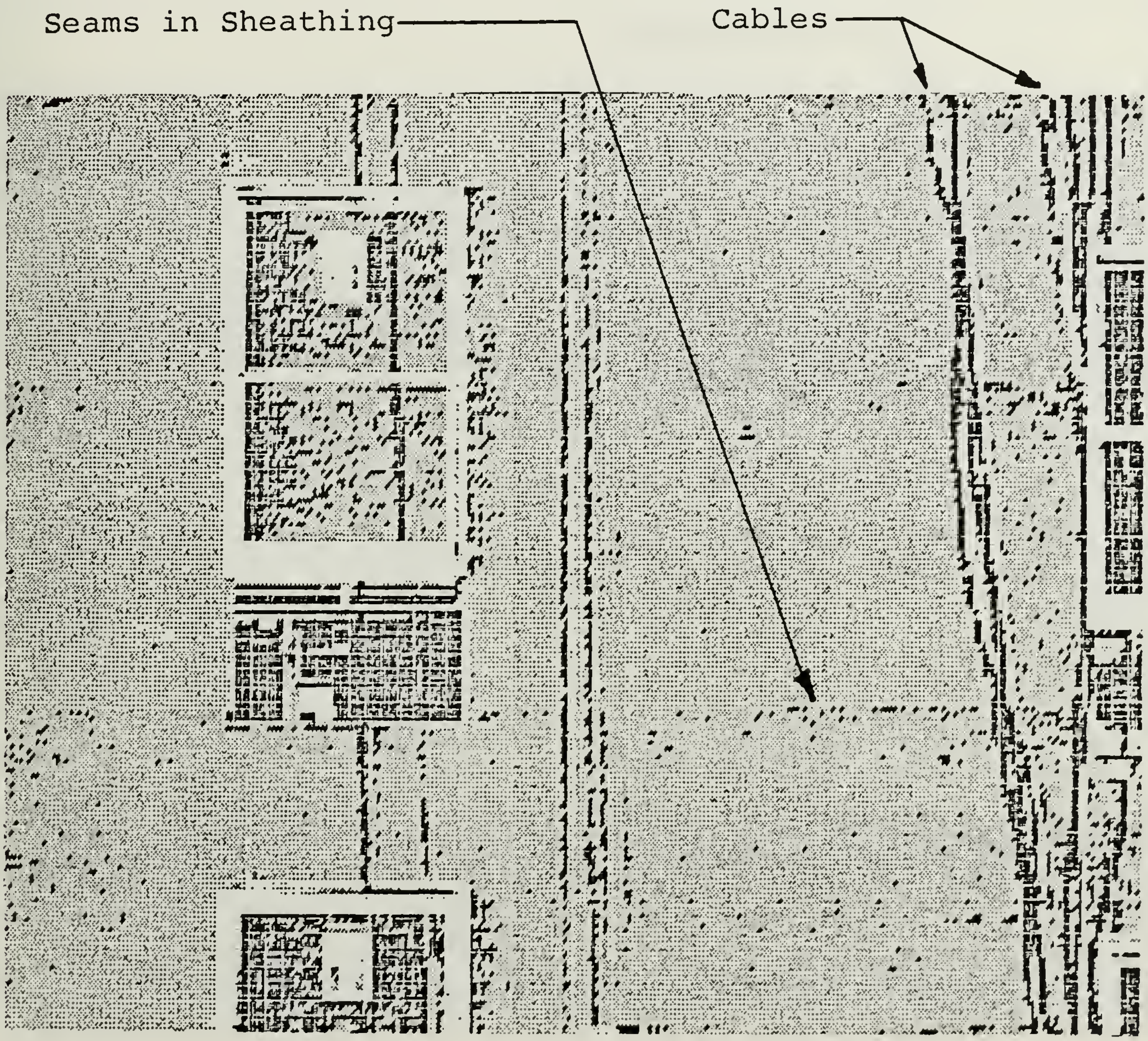
Wylie, C. R. and Barrett, L. C., Advanced Engineering Mathematics, McGraw-Hill Book Company, New York, 1982.

APPENDIX A

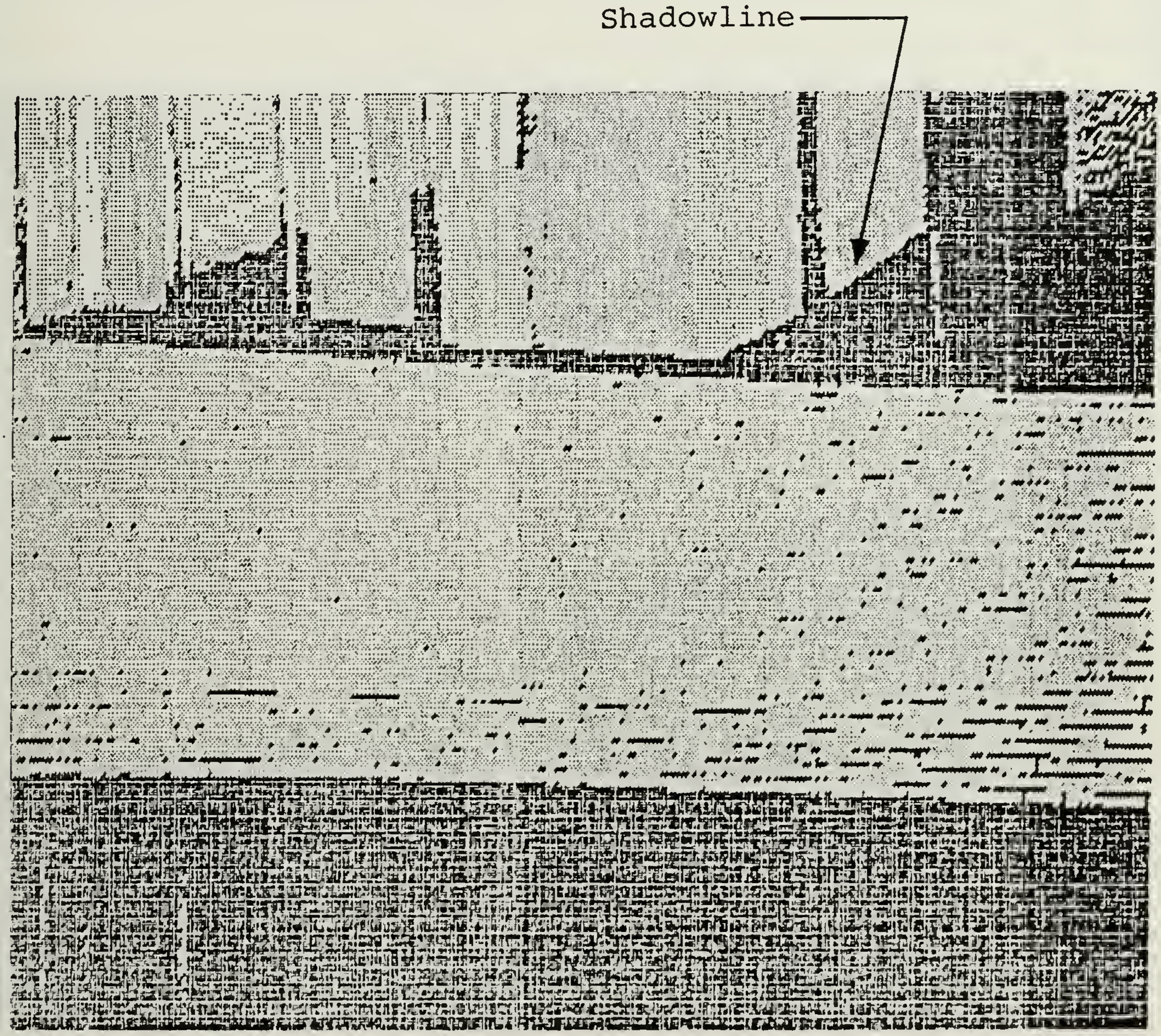
HARD COPY PCEYE IMAGES



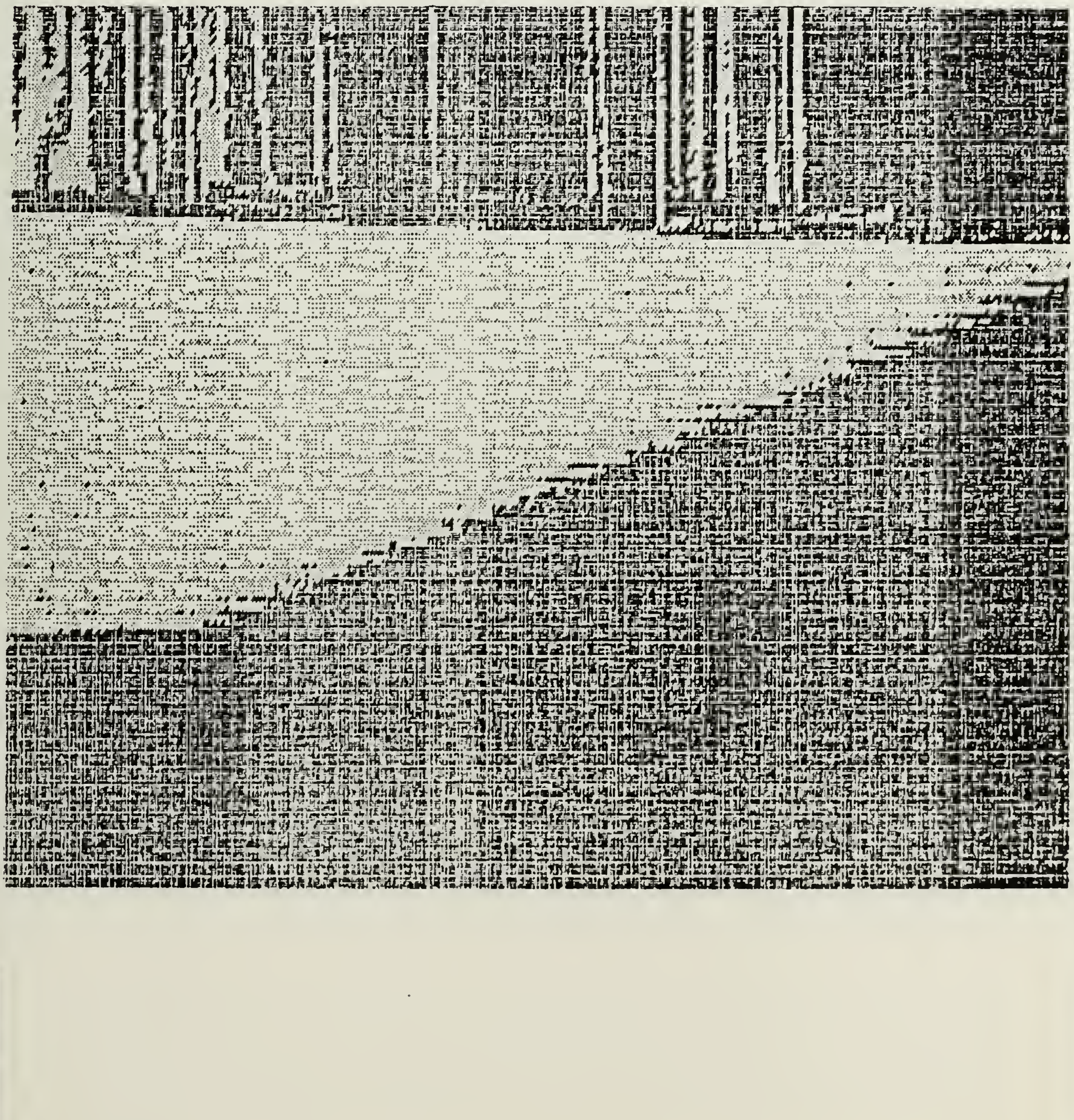
CLOSE-UP IMAGE OF ATHERTON HOTEL
WITH BARE STUDS
FIGURE A.1



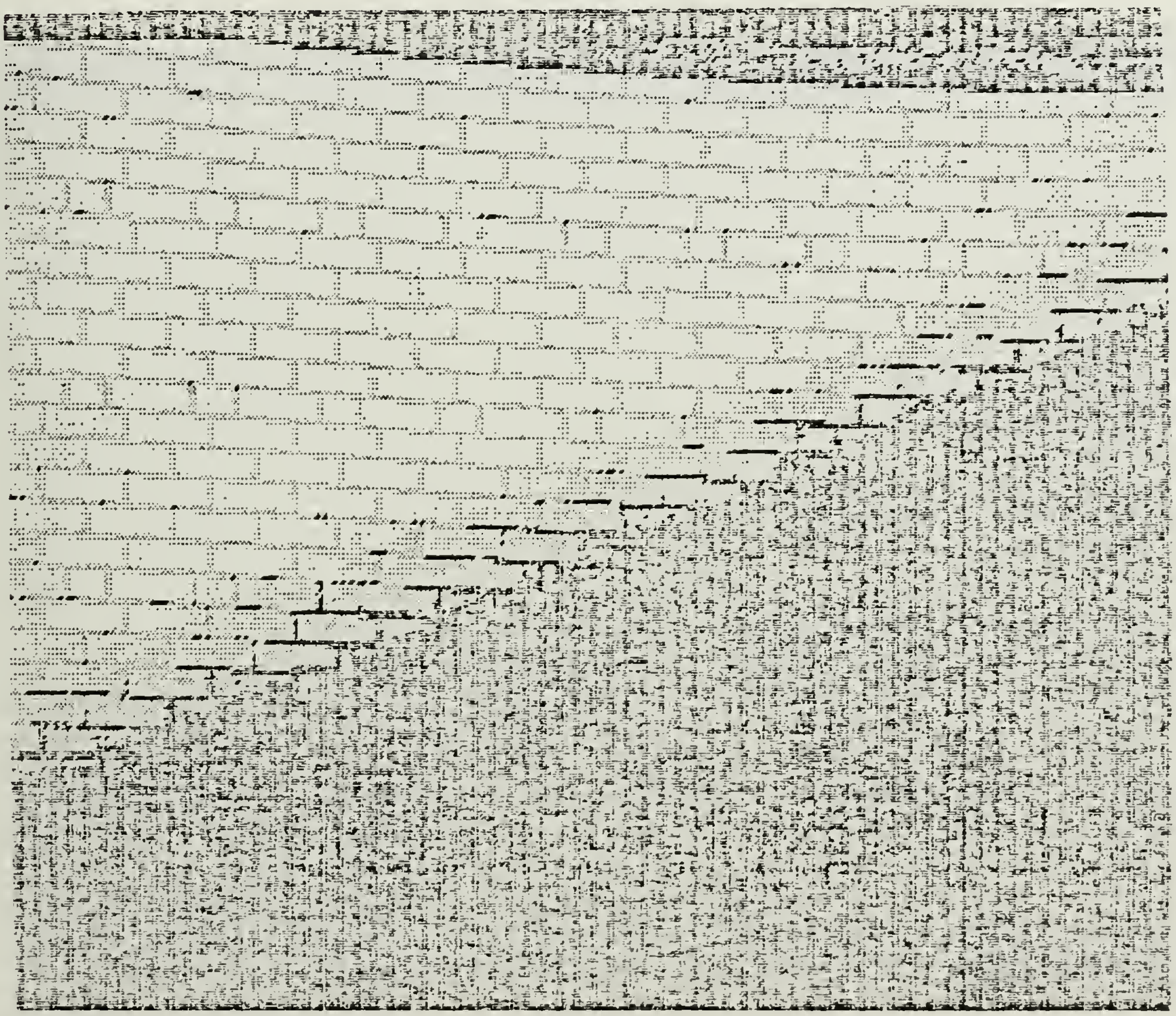
CLOSE-UP OF ATHERTON HOTEL
AFTER SHEATHING INSTALLED
FIGURE A.2



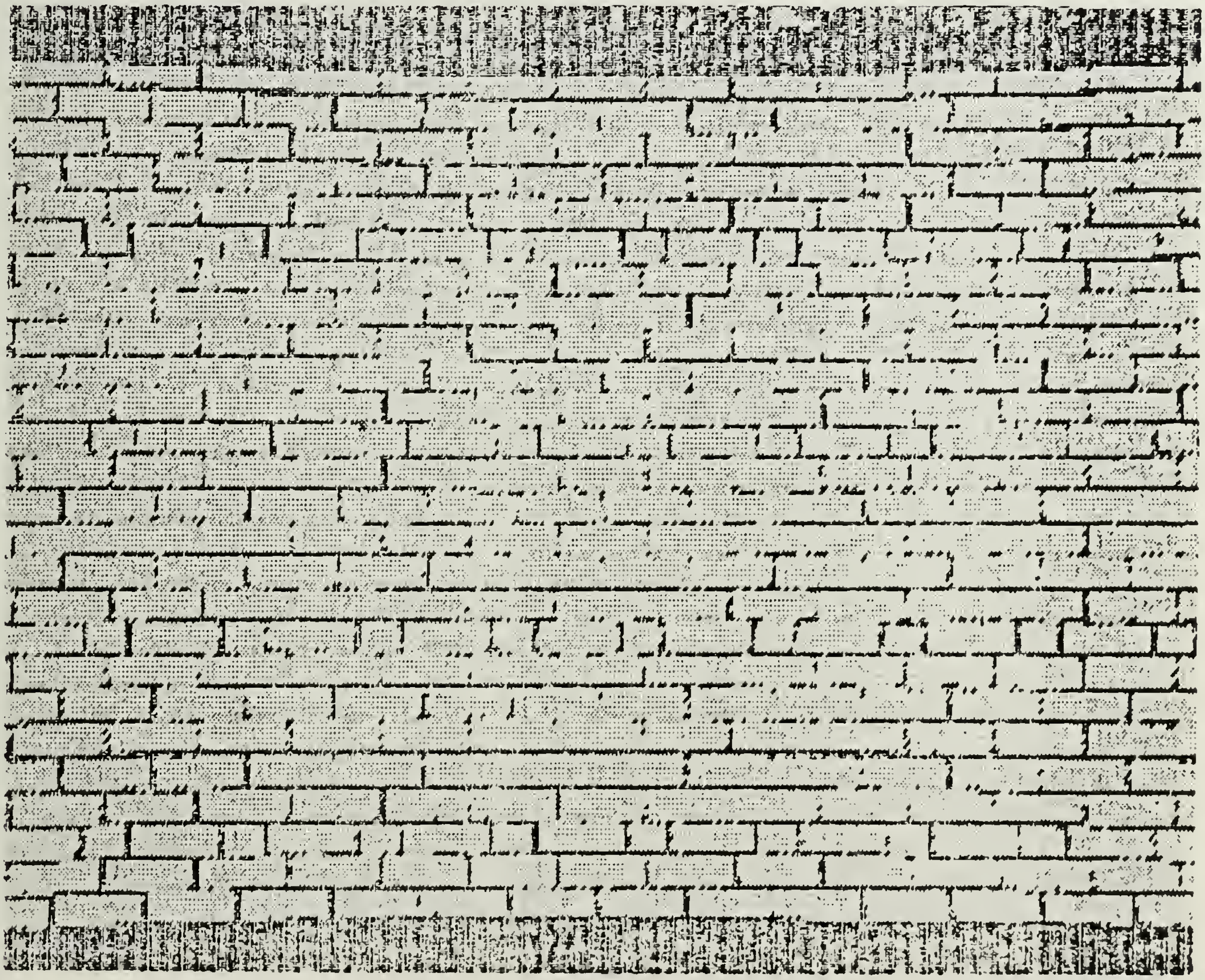
CLOSE-UP OF CENTRE COMMUNITY HOSPITAL
HIGHLIGHTING SHADOWLINE
FIGURE A.3



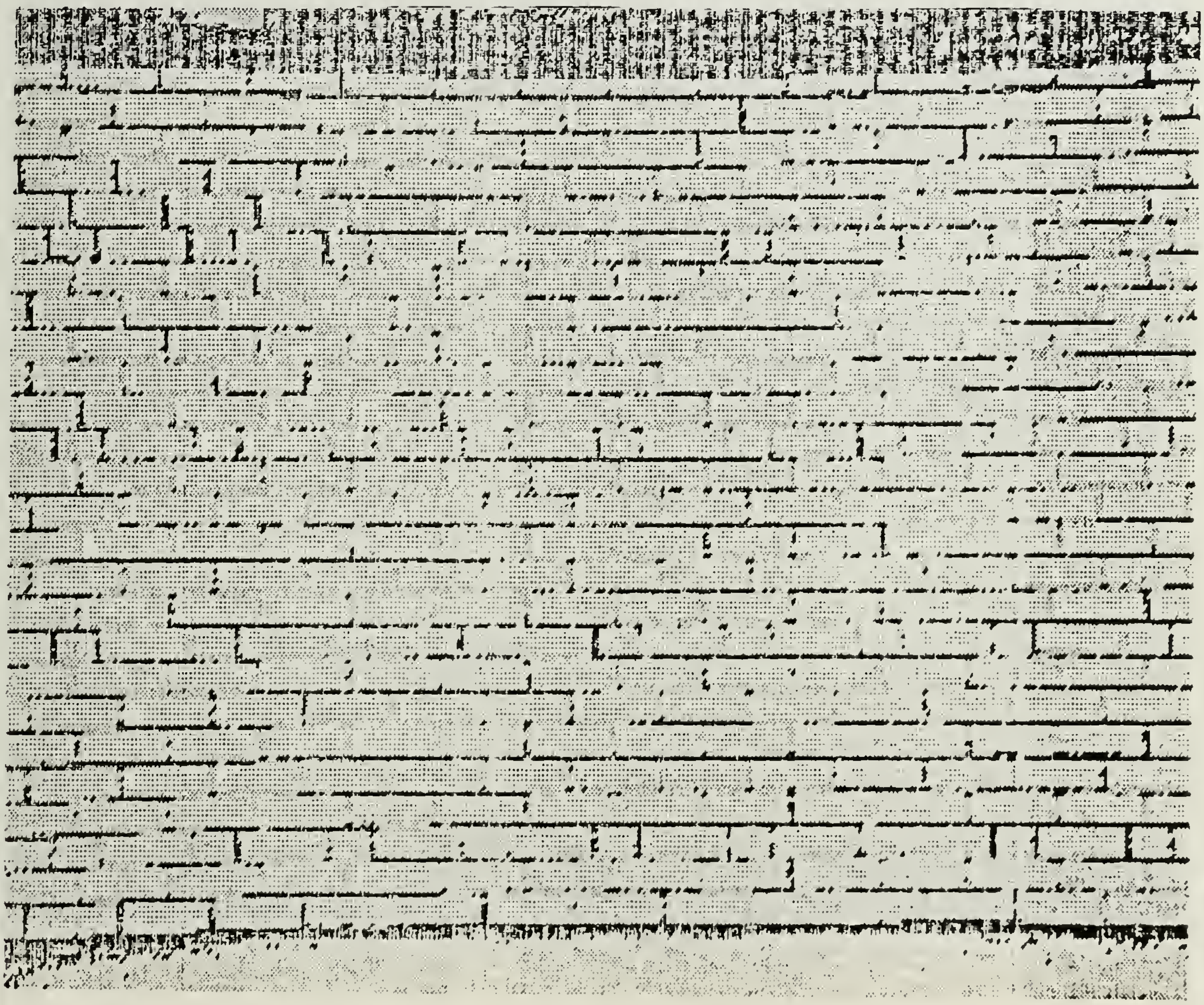
CLOSE-UP OF CENTRE COMMUNITY HOSPITAL
HIGHLIGHTING SHADOWLINE
FIGURE A.4



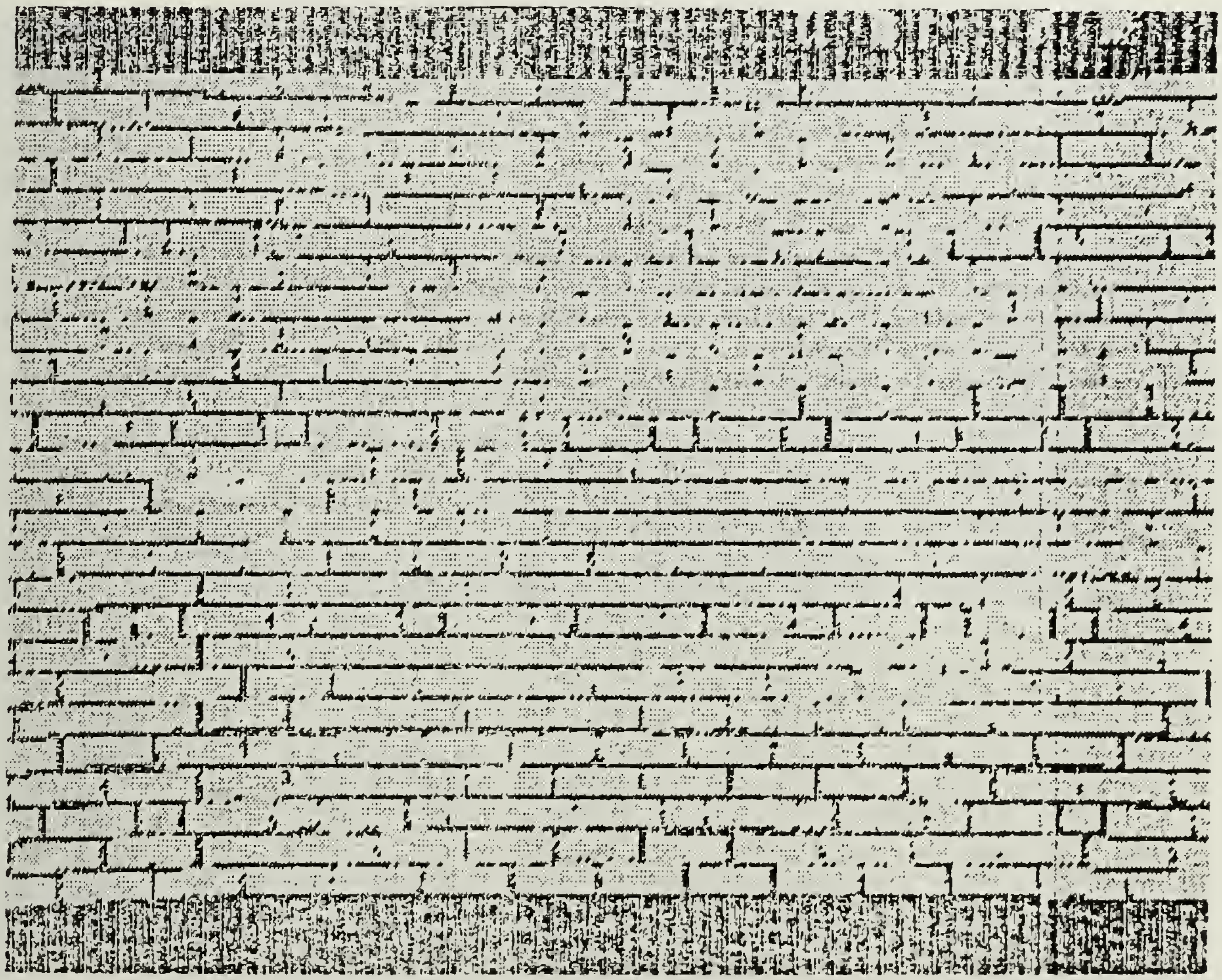
CLOSE-UP OF CENTRE COMMUNITY HOSPITAL
HIGHLIGHTING SHADOWLINE
FIGURE A.5



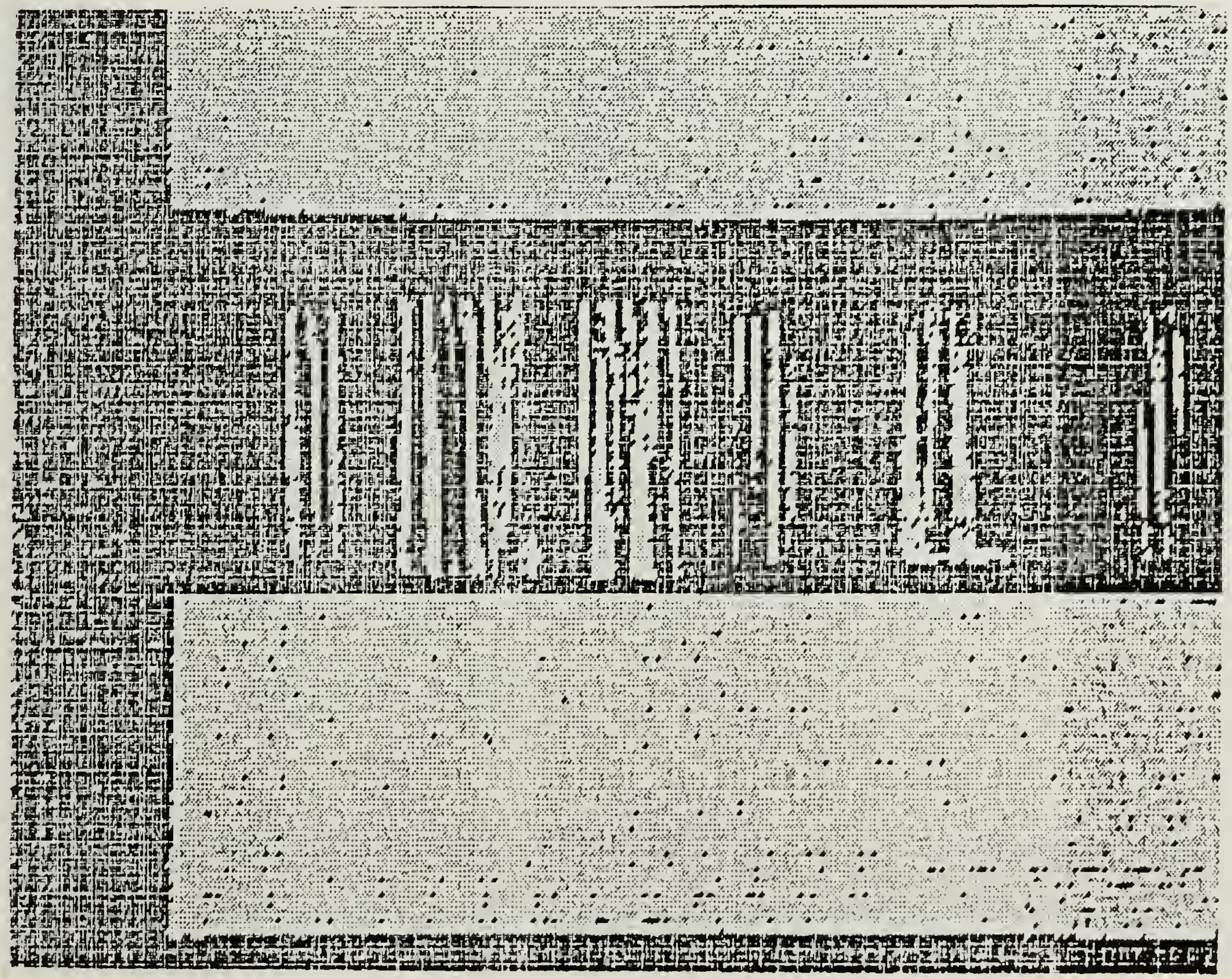
MORNING CLOSE-UP OF CENTRE COMMUNITY HOSPITAL.
NOTE CURVED EDGES AT TOP AND BOTTOM OF IMAGE.
FIGURE A.6



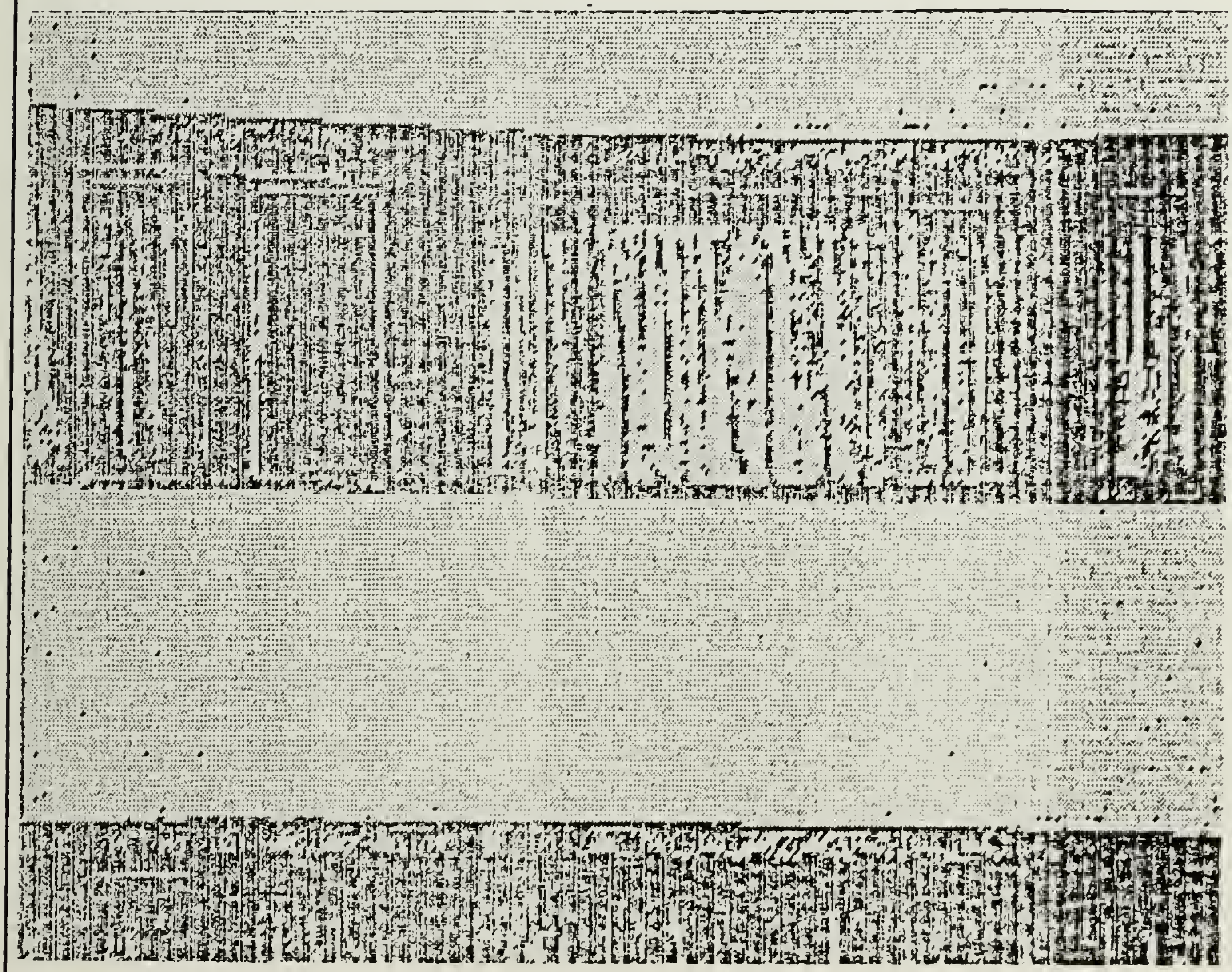
NOON CLOSE-UP OF CENTRE COMMUNITY HOSPITAL.
NOTE CURVED EDGES AT TOP AND BOTTOM OF IMAGE.
FIGURE A.7



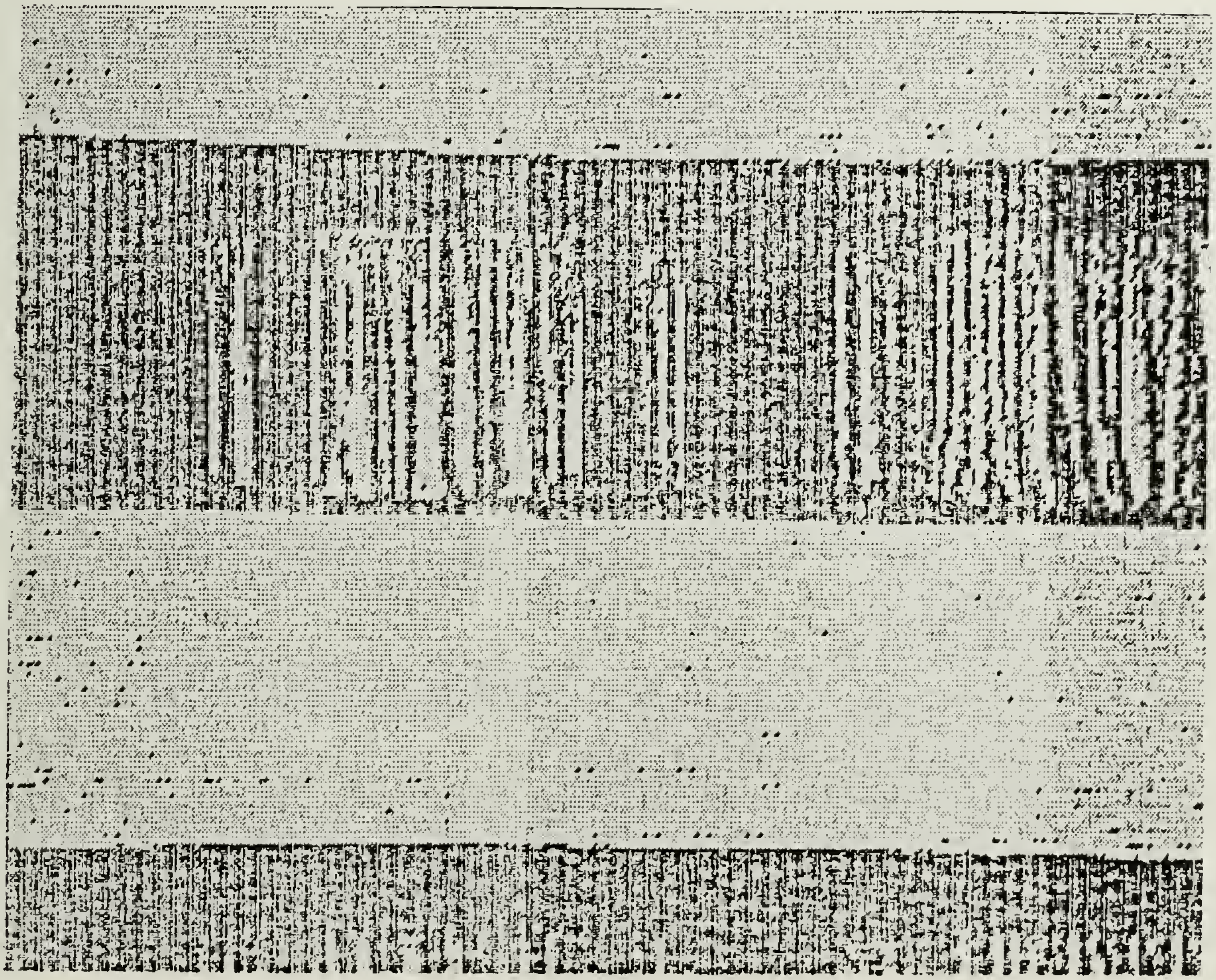
AFTERNOON CLOSE-UP OF CENTRE COMMUNITY HOSPITAL.
NOTE CURVED EDGES AT TOP AND BOTTOM OF IMAGE.
FIGURE A.8



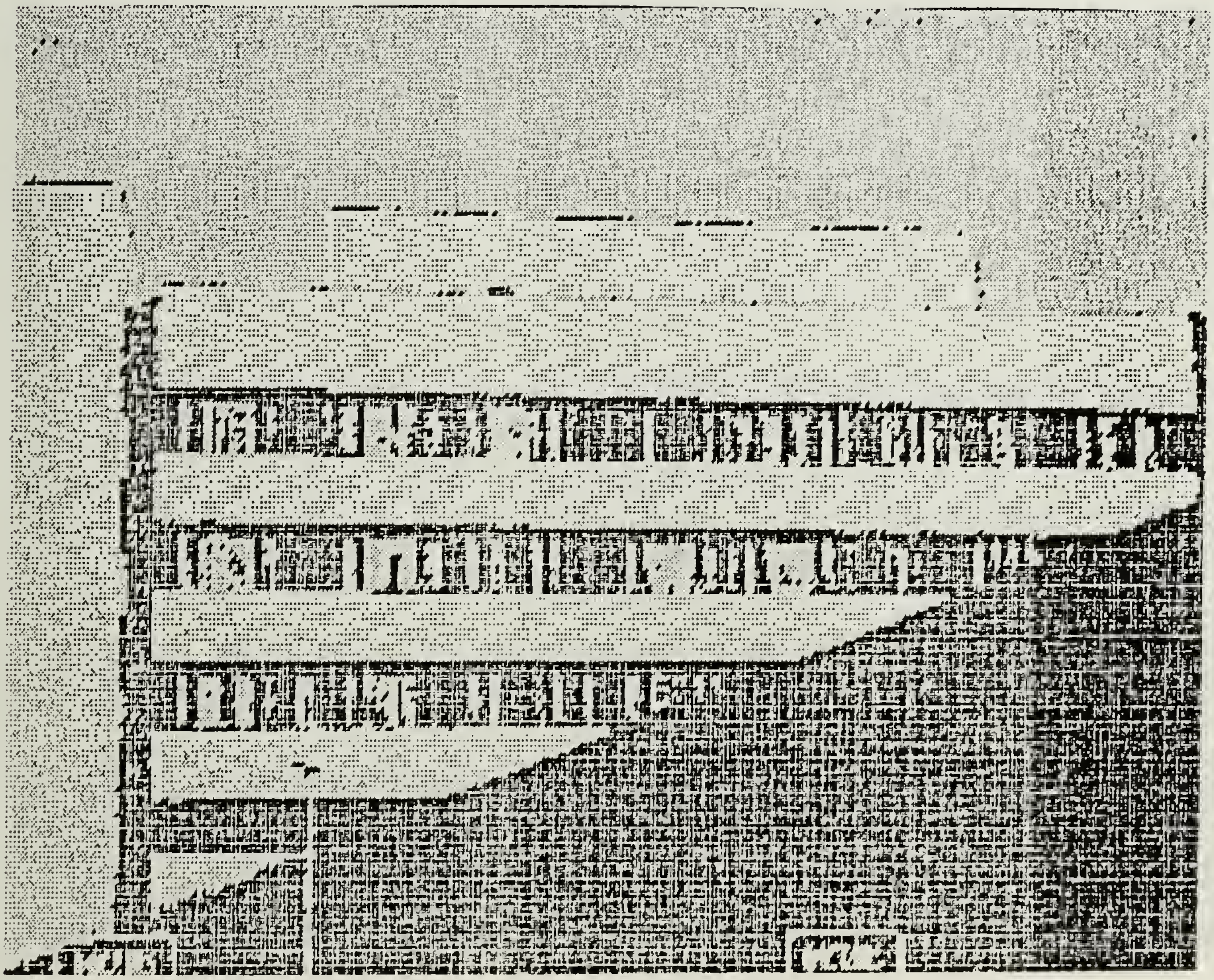
CENTRE COMMUNITY HOSPITAL
9 A.M., 300 FEET, FULL ZOOM
FIGURE A.9



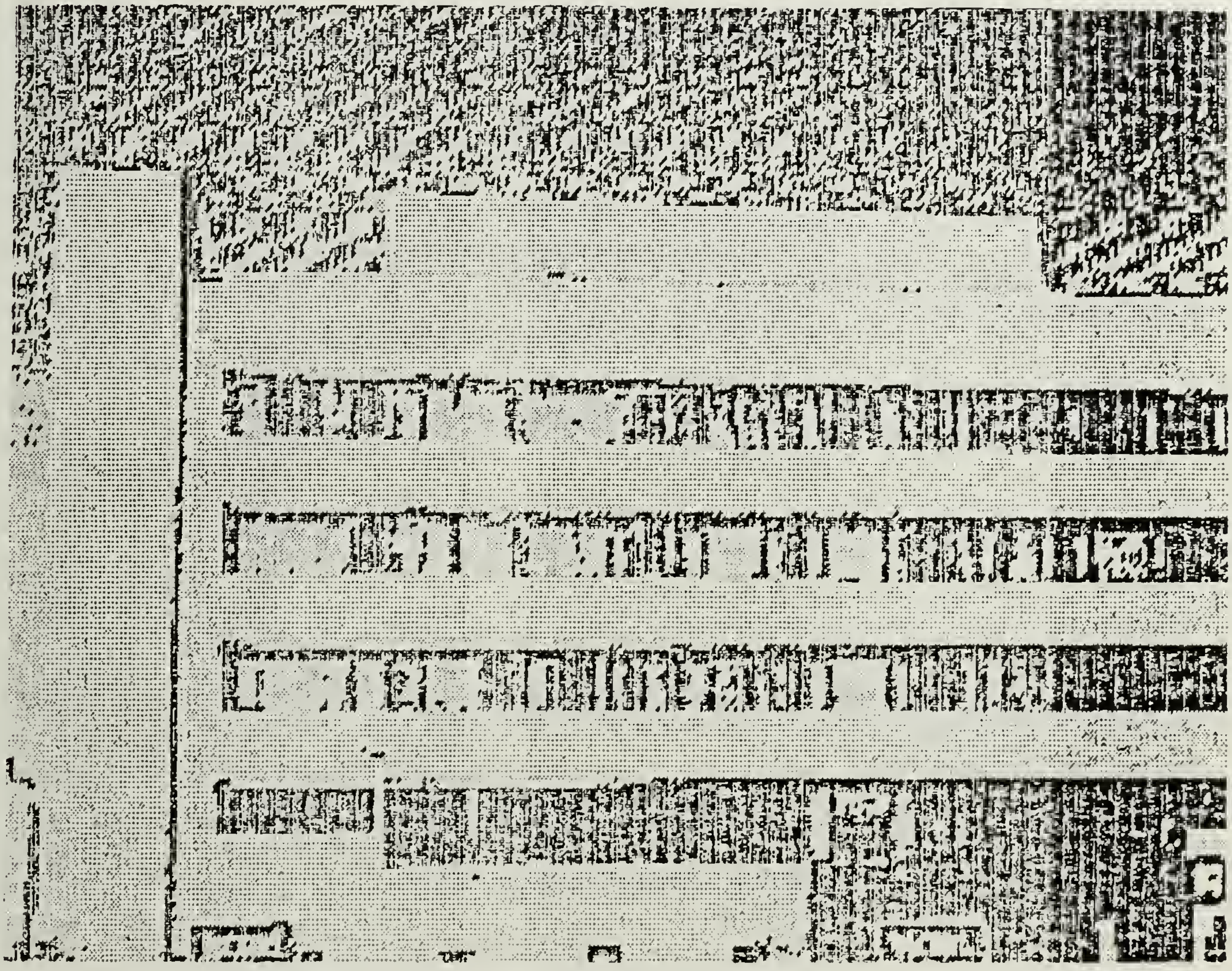
CENTRE COMMUNITY HOSPITAL
NOON, 300 FEET, FULL ZOOM
FIGURE A.10



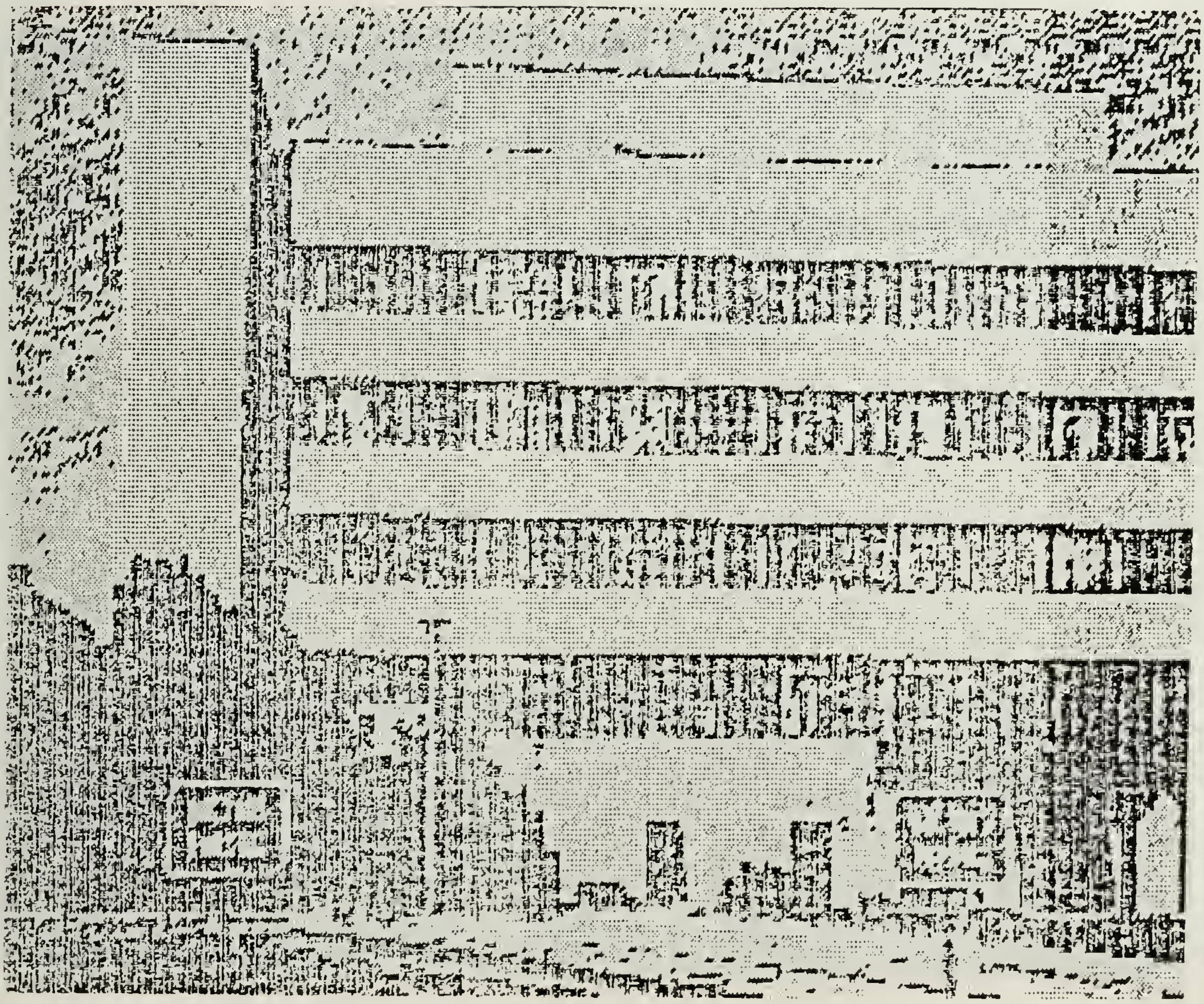
CENTRE COMMUNITY HOSPITAL
3 P.M., 300 FEET, FULL ZOOM
FIGURE A.11



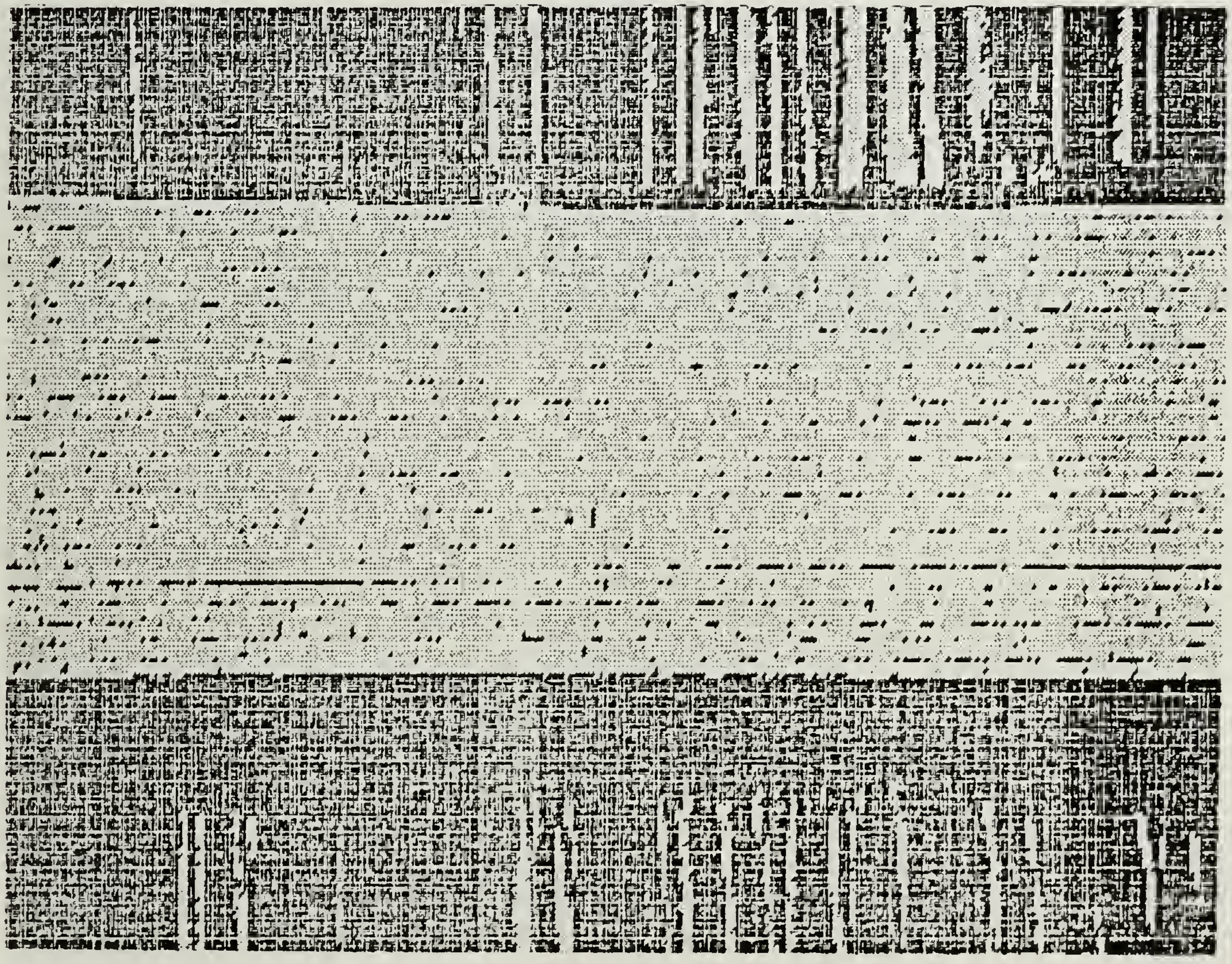
CENTRE COMMUNITY HOSPITAL
9 A.M., 200 FEET, FULL WIDE
FIGURE A.12



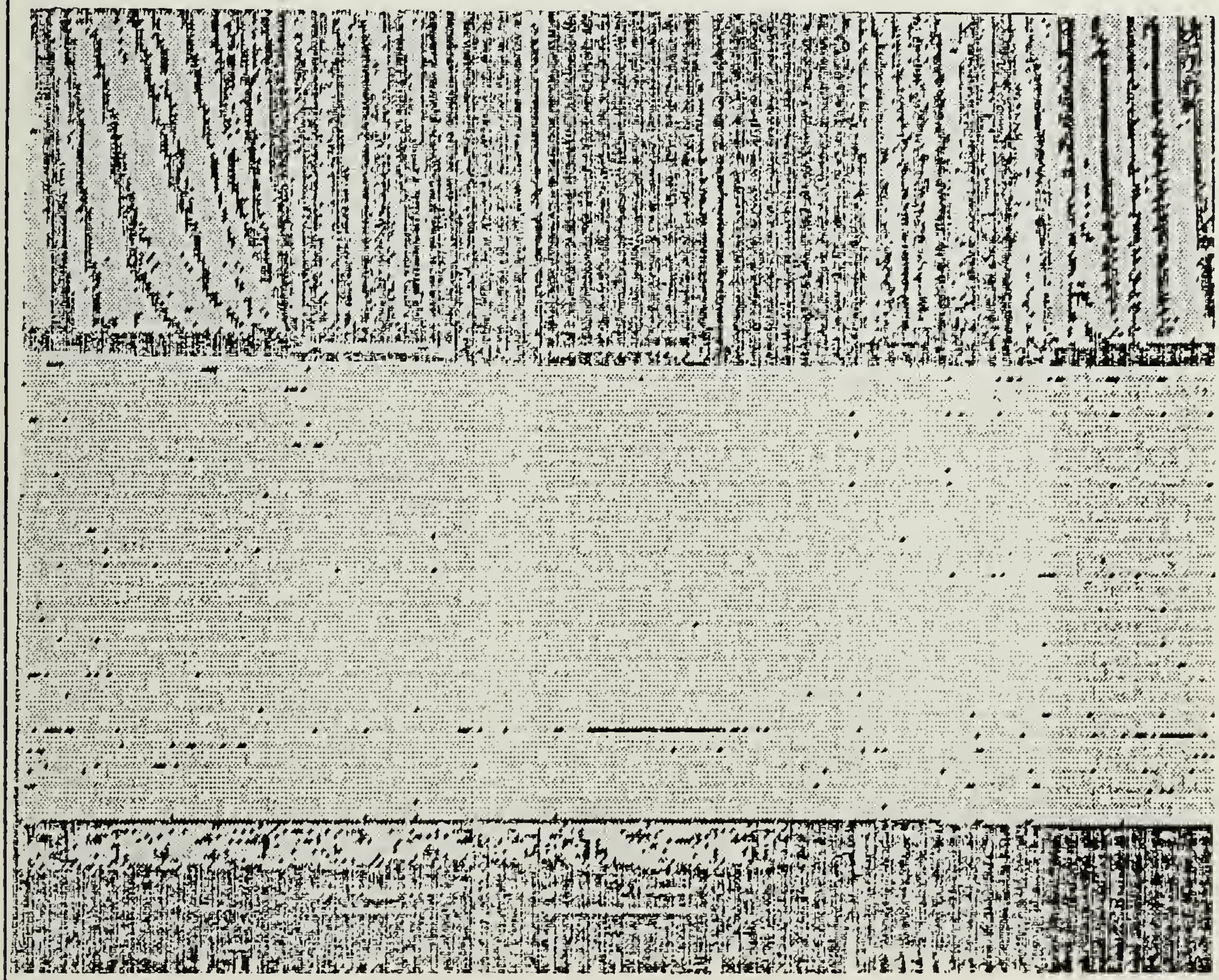
CENTRE COMMUNITY HOSPITAL
NOON, 200 FEET, FULL WIDE
FIGURE A.13



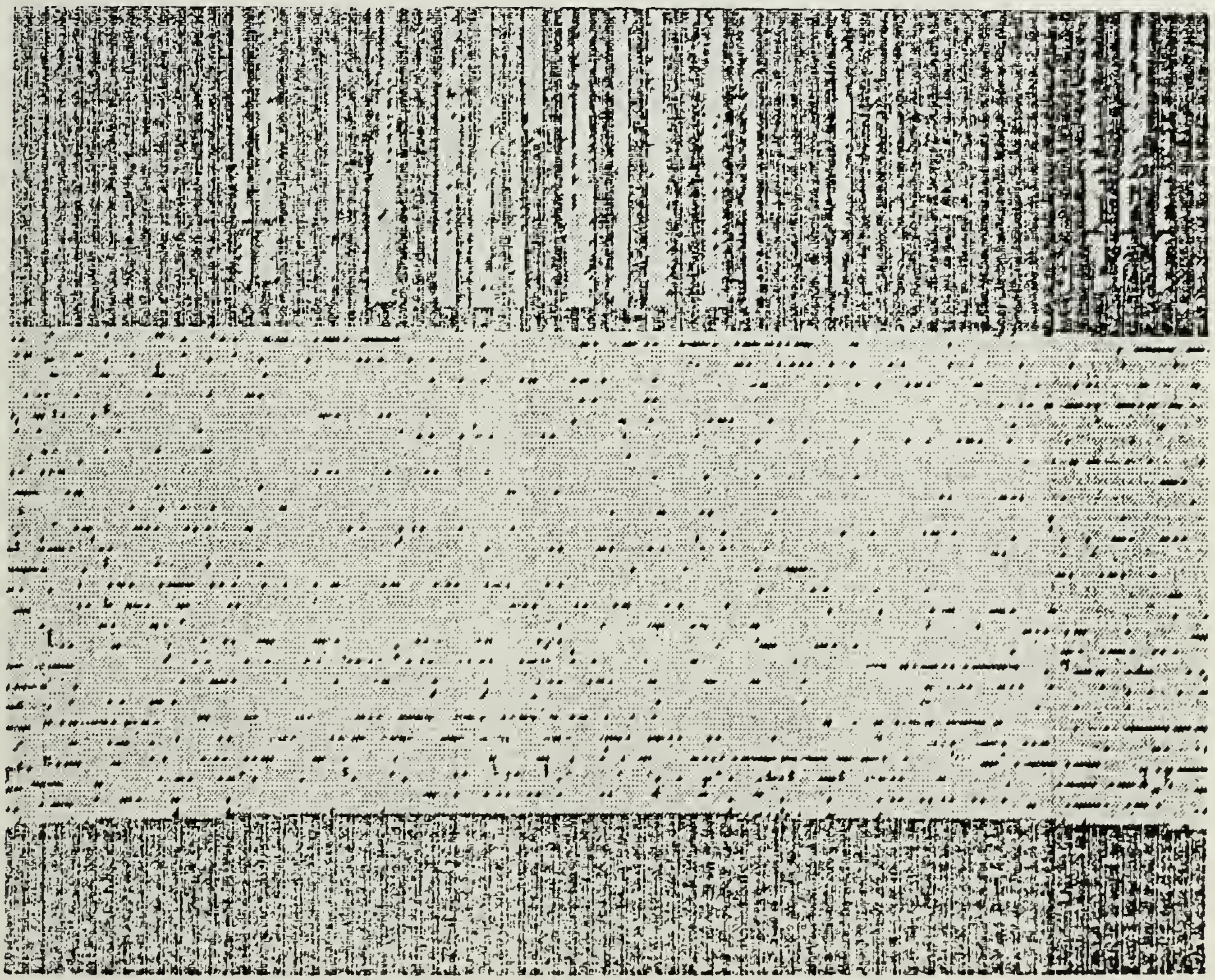
CENTRE COMMUNITY HOSPITAL
3 P.M., 200 FEET, FULL WIDE
FIGURE A.14



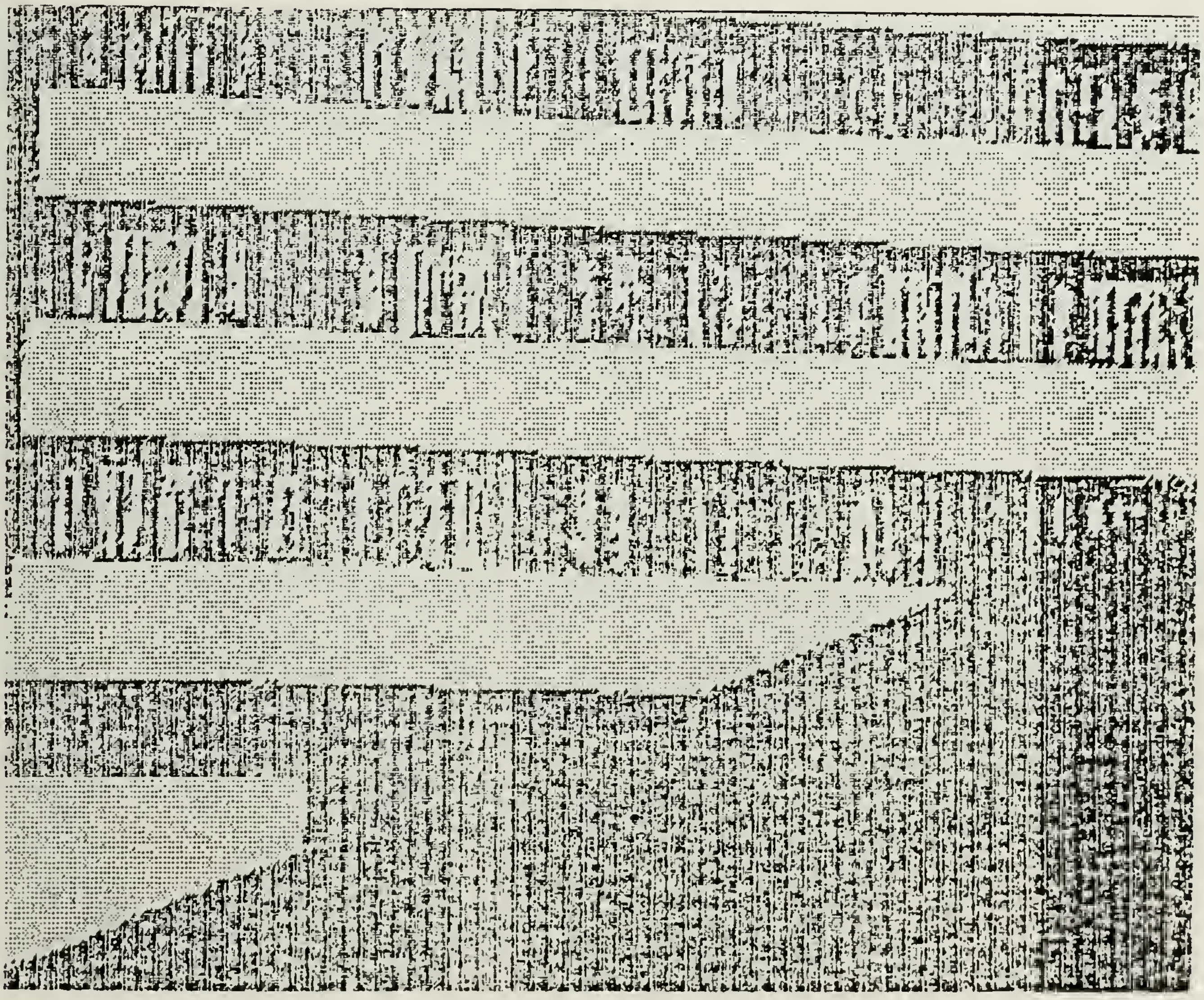
CENTRE COMMUNITY HOSPITAL
9 A.M., 200 FEET, FULL ZOOM
FIGURE A.15



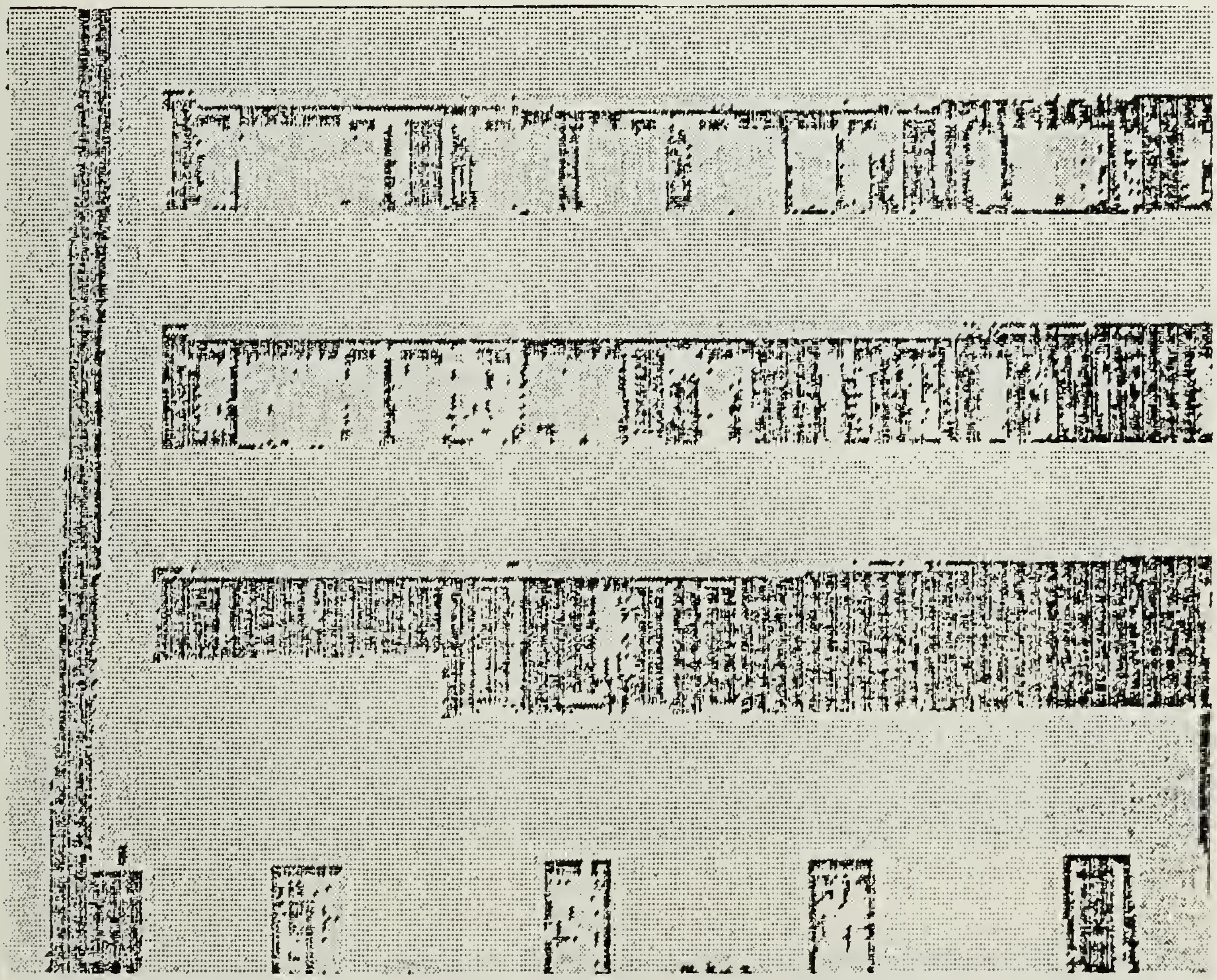
CENTRE COMMUNITY HOSPITAL
NOON, 200 FEET, FULL ZOOM
FIGURE A.16



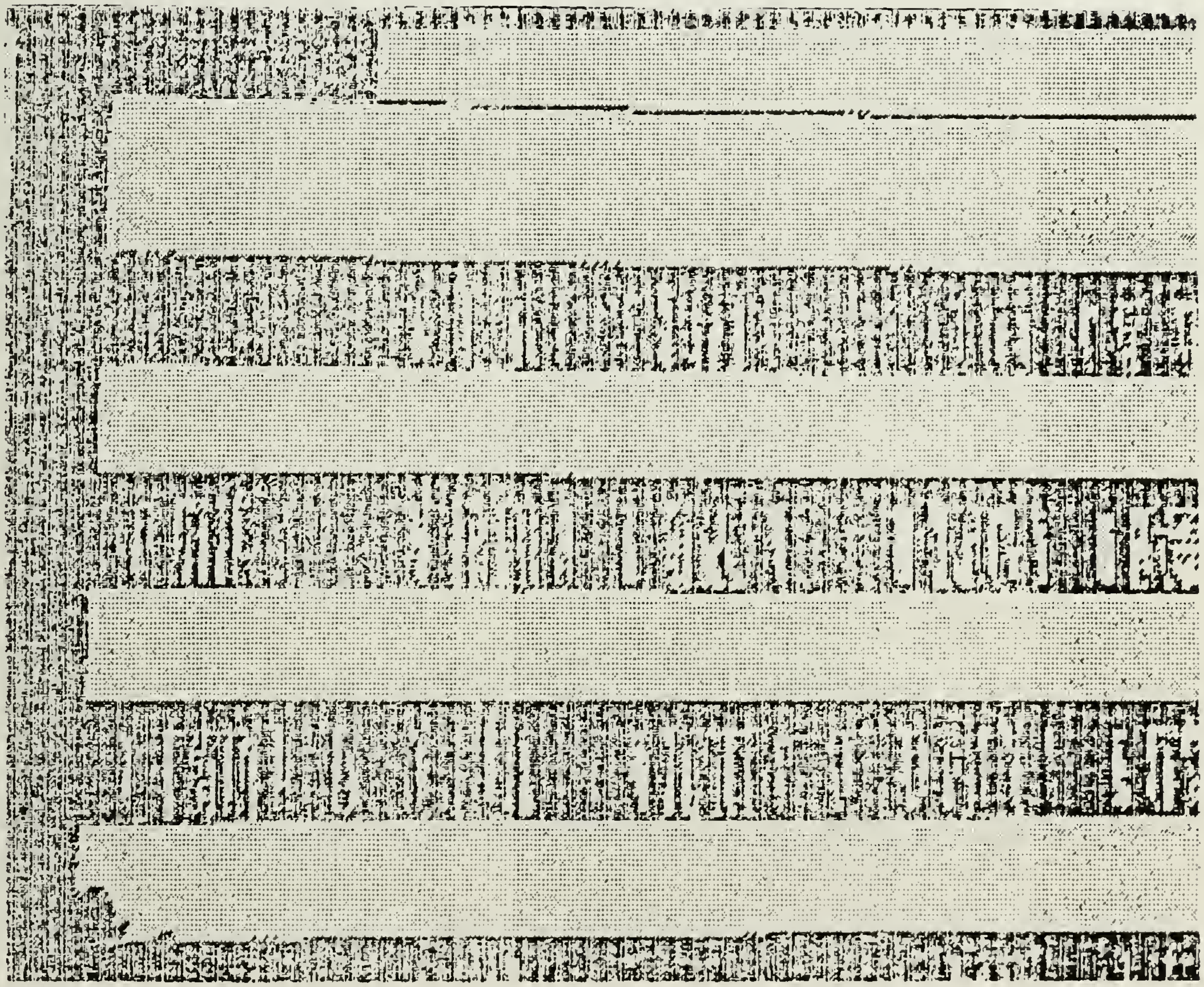
CENTRE COMMUNITY HOSPITAL
3 P.M., 200 FEET, FULL ZOOM
FIGURE A.17



CENTRE COMMUNITY HOSPITAL
9 A.M., 100 FEET, FULL WIDE
FIGURE A.18



CENTRE COMMUNITY HOSPITAL
NOON, 100 FEET, FULL WIDE
FIGURE A.19



CENTRE COMMUNITY HOSPITAL
3 P.M., 100 FEET, FULL WIDE
FIGURE A.20

APPENDIX B

DERIVATION OF PR_m

This appendix contains the derivation of the horizontal pixel ratio used in Chapter 4. The variables used in the derivation are as follows:

D_h = horizontal distance from camera to object
b = offset distance to reference object (perp. to D_h)
D_o = distance to reference object
D = distance to any point in the image
C_o = screen column of the centerline of fixed dimension
α_h = horizontal angle of the camera with respect to perpendicular from building centerline
σ_h = the horizontal angular field of view
δ = angular separation from reference point

All variables used in this derivation are illustrated in Figure 4.4.

$$D_o = \sqrt{(D_h^2 + b^2)}$$
$$D = \sqrt{(D_h^2 + (b+db)^2)} \qquad \text{(B-1)}$$

But $b = D_h \tan\alpha_{ho}$ and $db = D_h(\tan(\alpha_{ho}+\delta)-\tan\alpha_{ho})$; therefore, by substitution:

$$D_o = D_h\sqrt{(1 + \tan^2\alpha_{ho})} = D_h \sec\alpha_{ho}$$

and $$D = D_h\sqrt{(1 + \tan^2(\alpha_{ho} + \delta))} = D_h \sec(\alpha_{ho} + \delta). \qquad \text{(B-2)}$$

Since the vertical field of view is represented by

$$FOV = 2\ D\ \tan(\sigma_v/2) \qquad \text{(B-3)}$$

the ratio of the two fields of view shown above is

$$\text{FOV ratio} = \frac{FOV}{FOV_o} = \frac{2(D_h \sec(\alpha_{ho} + \delta))\tan(\sigma_v/2)}{2(D_h \sec\alpha_{ho})\tan(\sigma_v/2)} \qquad \text{(B-4)}$$

which simplifies to:

$$\text{FOV ratio} = \frac{\sec(\alpha_{ho} + \delta)}{\sec(\alpha_{ho})} = \frac{\cos(\alpha_{ho})}{\cos(\alpha_{ho} + \delta)} \qquad \text{(B-5)}$$

For the purpose of this derivation, σ_h has been chosen as the horizontal angle of equivalence, so that the angular separation equates on the screen to:

$$\delta = \frac{C - C_o}{360}\,\sigma_h \qquad \text{(B-6)}$$

The field of view ratio now becomes the Pixel Ratio by substitution for δ in the FOV ratio equation (B-5).

$$PR_h = \text{PIXEL RATIO} = \frac{\cos(\alpha_{ho})}{\cos\left(\alpha_{ho} + \dfrac{C - C_o}{360}\,\sigma_h\right)} \qquad \text{(B-7)}$$

The h subscript indicates that this pixel ratio adjusts for the distortion due to the differing distance in the horizontal direction.

APPENDIX C

BASIC PROGRAMS TO DETERMINE PIXEL RATIOS

```
100 '**********************************************************************
110 '**
120 '**
130 '**          PROGRAM TO DETERMINE MAXIMUM AND MINIMUM PRV
140 '**
150 '**
160 '**********************************************************************
170 PI = 4*ATN(1)
180 PRINT "ALPHAV";" SIGMAV","PRSUB1","PRSUB2","PRSUB3","PRSUB4"
190 PRINT
200 PRMAX = 0
210 PRMIN = 100
220 '45 DEGREES IS THE MAXIMUM VERTICAL ANGLE CONSIDERED
230 FOR ALPHAV = 0 TO 45 STEP 5
240  ALPHAVEE = ALPHAV * PI/180
250  PRINT ALPHAV
260   'CAMERA USED FOR RESEARCH HAS VERTICAL FOV OF 5-33 DEGREES
270   FOR SIGMAV = 5 TO 33 STEP 4
280    SIGMAVEE = SIGMAV * PI/180
290    'CASE 1:  REFERENCE AT TOP OF SCREEN, POINT OF INTEREST AT BOTTOM.
300    PRSUB1 = COS(ALPHAVEE + SIGMAVEE*.5)/COS(ALPHAVEE + SIGMAVEE*(-.5)
310    'CASE 2:  REFERENCE AT TOP OF SCREEN, POINT OF INTEREST AT MIDDLE.
320    PRSUB2 = COS(ALPHAVEE + SIGMAVEE*.5)/COS(ALPHAVEE)
330    'CASE 3:  REFERENCE AT MIDDLE OF SCREEN, POINT OF INTEREST AT TOP.
340    PRSUB3 = COS(ALPHAVEE)/COS(ALPHAVEE + SIGMAVEE * .5)
350    'CASE 4:  REFERENCE AT MIDDLE OF SCREEN, POINT OF INTEREST AT BOTT
360    PRSUB4 = COS(ALPHAVEE)/COS(ALPHAVEE + SIGMAVEE *(-.5))
370   PRINT SPC(8);SIGMAV,PRSUB1,PRSUB2,PRSUB3,PRSUB4
380  IF PRMAX < PRSUB1 THEN PRMAX = PRSUB1:MAX=SIGMAV:SUB=1
390  IF PRMAX < PRSUB2 THEN PRMAX = PRSUB2:MAX=SIGMAV:SUB=2
400  IF PRMAX < PRSUB3 THEN PRMAX = PRSUB3:MAX=SIGMAV:SUB=3
410  IF PRMAX < PRSUB4 THEN PRMAX = PRSUB4:MAX=SIGMAV:SUB=4
420  IF PRMIN > PRSUB1 THEN PRMIN = PRSUB1:MIN=SIGMAV:SU=1
430  IF PRMIN > PRSUB2 THEN PRMIN = PRSUB2:MIN=SIGMAV:SU=2
440  IF PRMIN > PRSUB3 THEN PRMIN = PRSUB3:MIN=SIGMAV:SU=3
450  IF PRMIN > PRSUB4 THEN PRMIN = PRSUB4:MIN=SIGMAV:SU=4
460   NEXT SIGMAV
470 NEXT ALPHAV
480 END
```

ALPHAV	SIGMAV	PRSUB1	PRSUB2	PRSUB3	PRSUB4
0					
	5	1	.9990483	1.000953	1.000953
	9	1	.9969174	1.003092	1.003092
	13	1	.9935719	1.00647	1.00647
	17	1	.9890158	1.011106	1.011106
	21	1	.983255	1.01703	1.01703
	25	1	.9762961	1.02428	1.02428
	29	1	.9681476	1.0329	1.0329
	33	1	.9580198	1.042949	1.042949
5					
	5	.9923894	.9952321	1.004791	.9971437
	9	.9863231	.9900531	1.010047	.9962326
	13	.9802606	.9836678	1.016603	.9965361
	17	.9741869	.9760842	1.024502	.9980562
	21	.9680872	.9673113	1.033793	1.000802
	25	.9619465	.95736	1.044539	1.004791
	29	.9557491	.9462422	1.056812	1.010047
	33	.9494787	.9339716	1.070696	1.016603
10					
	5	.9847204	.991357	1.008718	.9933055
	9	.9726254	.9830829	1.017208	.9893625
	13	.9606114	.9736111	1.027104	.9866479
	17	.9486487	.9629531	1.038472	.9851452
	21	.9367079	.9511219	1.05139	.9848452
	25	.9247597	.9381319	1.065948	.9857459
	29	.912775	.9239989	1.082252	.9878529
	33	.9007243	.9087401	1.100425	.9911792
15					
	5	.9768727	.9873604	1.012801	.989378
	9	.9586949	.9758944	1.024701	.9823757
	13	.9407509	.9632392	1.038164	.9766536
	17	.9229931	.9494104	1.053285	.9721751
	21	.9053762	.9344251	1.070177	.9689126
	25	.8878558	.9183012	1.088967	.966846
	29	.8703888	.9010585	1.109806	.9659625
	33	.8529324	.8827181	1.132865	.9662568
20					
	5	.9687146	.9831721	1.017116	.9852951
	9	.9443052	.9683604	1.032673	.9751588
	13	.920364	.9523691	1.050013	.9663942
	17	.8968211	.9352176	1.06927	.9589433
	21	.8736101	.9169265	1.0906	.9527591
	25	.8506689	.8975183	1.114183	.9478012
	29	.8279378	.8770167	1.140229	.9440389
	33	.805359	.8554465	1.16893	.9414486

25					
	5	.9600936	.9787081	1.021755	.9809804
	9	.9291998	.9603312	1.041307	.9675827
	13	.8991024	.9407842	1.062943	.9556946
	17	.8697002	.9200911	1.086849	.9452328
	21	.8409002	.8982771	1.113242	.9361256
	25	.8126155	.8753684	1.142376	.9283125
	29	.7847654	.8513934	1.174545	.9217424
	33	.7572737	.8263811	1.210095	.9163734
30					
	5	.9508243	.9738646	1.026837	.9763415
	9	.9130729	.9516189	1.050841	.9594944
	13	.8765587	.9282139	1.077338	.9443498
	17	.8411365	.9036781	1.106589	.9307922
	21	.8066756	.8780411	1.138899	.9187219
	25	.7730567	.8513344	1.174627	.9080529
	29	.7401701	.8235907	1.214195	.8987111
	33	.7079143	.7948433	1.25811	.8906337
35					
	5	.9406704	.9685056	1.032519	.9712596
	9	.8955412	.9419796	1.061594	.9507014
	13	.8522316	.9143061	1.093726	.9321076
	17	.8105335	.8855185	1.129282	.9153208
	21	.7702628	.8556522	1.168699	.9002054
	25	.7312535	.8247433	1.212499	.8866438
	29	.6933568	.7928296	1.261305	.8745344
	33	.6564368	.75995	1.315876	.8637896
40					
	5	.9293178	.9624473	1.039018	.9655778
	9	.8761046	.9310824	1.074019	.9409528
	13	.8254781	.8985831	1.112868	.9186442
	17	.7771391	.8649891	1.156084	.8984381
	21	.7308255	.8303412	1.204324	.8801509
	25	.6863066	.7946817	1.258366	.8636246
	29	.6433773	.7580538	1.319168	.8487224
	33	.6018547	.7205025	1.38792	.8353263
45					
	5	.9163312	.9554288	1.04665	.9590784
	9	.8540806	.9184582	1.088781	.9299069
	13	.7954358	.8803686	1.135888	.9035259
	17	.739961	.8412065	1.188769	.8796426
	21	.687231	.8010194	1.248409	.8580079
	25	.6370702	.7598563	1.316038	.8384087
	29	.589045	.7177676	1.392209	.8206625
	33	.5429556	.6748044	1.481911	.804612

```
100 '***********************************************************************
110 '**                                                                   **
120 '**                                                                   **
130 '**          PROGRAM TO DETERMINE MAXIMUM AND MINIMUM PRH             **
140 '**                                                                   **
150 '**                                                                   **
160 '***********************************************************************
170 PI = 4*ATN(1)
180 PRINT "ALPHAH";" SIGMAH","PRSUB1","PRSUB2","PRSUB3","PRSUB4"
190 PRINT
200 PRMAX = 0
210 PRMIN = 100
220 '45 DEGREES OFF CENTER WOULD BE THE MAXIMUM CONSIDERED
230 FOR ALPHAH = 0 TO 45 STEP 5
240  ALPHAHEE = ALPHAH * PI/180
250 PRINT ALPHAH
260     'THE CAMERA USED FOR THIS RESEARCH HAD HORIZ FOV OF 8-44 DEGREES
270    FOR SIGMAH = 8 TO 44 STEP 4
280     SIGMAHEE = SIGMAH * PI/180
290     'CASE 1:  REFERENCE AT FAR RIGHT, POINT OF INTEREST AT FAR LEFT.
300     PRSUB1 = COS(ALPHAHEE + SIGMAHEE*.5)/COS(ALPHAHEE + SIGMAHEE*(-.5)
310     'CASE 2:  REFERENCE AT FAR RIGHT, POINT OF INTEREST AT CENTER.
320     PRSUB2 = COS(ALPHAHEE + SIGMAHEE*.5)/COS(ALPHAHEE)
330     'CASE 3:  REFERENCE AT CENTER, POINT OF INTEREST AT FAR RIGHT.
340     PRSUB3 = COS(ALPHAHEE)/COS(ALPHAHEE + SIGMAHEE * .5)
350     'CASE 4:  REFERENCE AT CENTER, POINT OF INTEREST AT FAR LEFT.
360     PRSUB4 = COS(ALPHAHEE)/COS(ALPHAHEE + SIGMAHEE *(-.5))
370  PRINT SPC(8);SIGMAH,PRSUB1,PRSUB2,PRSUB3,PRSUB4
380  IF PRMAX < PRSUB1 THEN PRMAX = PRSUB1:MAX=SIGMAH:SUB=1
390  IF PRMAX < PRSUB2 THEN PRMAX = PRSUB2:MAX=SIGMAH:SUB=2
400  IF PRMAX < PRSUB3 THEN PRMAX = PRSUB3:MAX=SIGMAH:SUB=3
410  IF PRMAX < PRSUB4 THEN PRMAX = PRSUB4:MAX=SIGMAH:SUB=4
420  IF PRMIN > PRSUB1 THEN PRMIN = PRSUB1:MIN=SIGMAH:SU=1
430  IF PRMIN > PRSUB2 THEN PRMIN = PRSUB2:MIN=SIGMAH:SU=2
440  IF PRMIN > PRSUB3 THEN PRMIN = PRSUB3:MIN=SIGMAH:SU=3
450  IF PRMIN > PRSUB4 THEN PRMIN = PRSUB4:MIN=SIGMAH:SU=4
460   NEXT SIGMAH
470 PRMAX=0
480 PRMIN=100
490 MAX=0
500 MIN=0
510 SUB=0
520 SU=0
530 NEXT ALPHAH
540 END
```

25					
	8	.9368444	.9650361	1.036231	.970787
	12	.9065578	.9457794	1.057329	.9585298
	16	.8769908	.9253704	1.080648	.9477185
	20	.8480486	.9038342	1.106398	.9382789
	24	.8196428	.8811968	1.13482	.9301474
	28	.7916916	.8574857	1.1662	.9232709
	32	.764117	.8327298	1.20087	.9176049
	36	.736846	.8069595	1.23922	.913114
	40	.7098078	.7802059	1.281713	.9097698
	44	.6829342	.7525018	1.328901	.9075515
30					
	8	.9223888	.9572901	1.044616	.9635416
	12	.8855794	.9341723	1.070466	.9479829
	16	.8498969	.9099163	1.099002	.9340384
	20	.8152073	.8845518	1.130516	.9216049
	24	.7813886	.8581096	1.165352	.910593
	28	.7483288	.830622	1.203917	.9009258
	32	.7159244	.8021225	1.246693	.8925376
	36	.6840794	.7726454	1.294255	.885373
	40	.6527036	.7422271	1.347297	.8793853
	44	.6217119	.7109046	1.406659	.8745363
35					
	8	.9066444	.9487201	1.054052	.9556501
	12	.8629002	.9213302	1.085387	.9365808
	16	.8208175	.892818	1.120049	.9193559
	20	.7802059	.8632179	1.158456	.9038343
	24	.740896	.8325663	1.201106	.8898943
	28	.7027343	.8009002	1.248595	.8774306
	32	.6655822	.7682583	1.301646	.8663521
	36	.6293129	.7346804	1.361136	.8565805
	40	.59381	.7002075	1.428148	.8480486
	44	.5589652	.6648815	1.504028	.8406991
40					
	8	.8891528	.9390315	1.064927	.9468829
	12	.8379094	.9068122	1.102764	.9240164
	16	.7890243	.873488	1.144836	.9033029
	20	.742227	.8390996	1.191754	.8845518
	24	.6972796	.803689	1.244263	.8675988
	28	.653971	.7672092	1.303273	.8523024
	32	.6121128	.7299745	1.369911	.8385399
	36	.5715363	.6917605	1.445587	.8262054
	40	.5320888	.6527036	1.532089	.8152073
	44	.4936316	.6128516	1.631716	.8054666
45					
	8	.8692867	.9270076	1.07781	.9369256
	12	.809784	.8899934	1.123604	.9098763
	16	.753554	.8510949	1.174957	.8853936
	20	.7002075	.8111596	1.232803	.8632179
	24	.6494076	.770236	1.298304	.8431281
	28	.6008606	.7283738	1.372921	.8249343
	32	.5543089	.6856243	1.458525	.8084733
	36	.5095254	.6420396	1.557537	.7936045
	40	.4663076	.5976724	1.673157	.7802059
	44	.4244748	.5525773	1.809702	.7681727

ALPHAH	SIGMAH	PRSUB1	PRSUB2	PRSUB3	PRSUB4
0					
	8	1	.9975641	1.002442	1.002442
	12	1	.9945219	1.005508	1.005508
	16	1	.9902681	1.009828	1.009828
	20	1	.9848078	1.015427	1.015427
	24	1	.9781476	1.022341	1.022341
	28	1	.9702958	1.030614	1.030614
	32	1	.9612617	1.040299	1.040299
	36	1	.9510566	1.051462	1.051462
	40	1	.9396927	1.064178	1.064178
	44	1	.9271839	1.078535	1.078535
5					
	8	.9878309	.9914612	1.008612	.9963464
	12	.9817767	.9853768	1.01484	.9963465
	16	.9757072	.978092	1.022399	.9975618
	20	.9696154	.9696154	1.031337	1
	24	.9634865	.9599578	1.041713	1.003676
	28	.9573045	.9491303	1.053596	1.008612
	32	.9510541	.9371466	1.067069	1.01484
	36	.9447179	.924021	1.082226	1.022399
	40	.9382789	.9097698	1.099179	1.031337
	44	.9317181	.89441	1.118056	1.041713
10					
	8	.9756405	.9852642	1.014956	.9902324
	12	.9636089	.9760908	1.024495	.9872124
	16	.9516362	.9657282	1.035488	.985408
	20	.9396927	.954189	1.04801	.9848078
	24	.9277491	.9414872	1.062149	.985408
	28	.9157761	.9276384	1.078006	.9872124
	32	.9037449	.9126594	1.095699	.9902324
	36	.8916248	.8965686	1.115364	.994486
	40	.8793852	.8793852	1.137158	1
	44	.866994	.8611306	1.161264	1.006809
15					
	8	.9632155	.9788729	1.021583	.9840048
	12	.9452176	.9665137	1.034647	.9779661
	16	.9274176	.9529768	1.049344	.9731797
	20	.9097697	.9382789	1.065781	.9696154
	24	.8922292	.9224379	1.084084	.9672513
	28	.8747529	.905473	1.104395	.9660729
	32	.8572978	.8874049	1.126881	.9660729
	36	.8398214	.8682556	1.151735	.9672514
	40	.8222811	.8480486	1.179178	.9696154
	44	.8046331	.8268083	1.20947	.9731797
20					
	8	.9503608	.9721747	1.028622	.9775618
	12	.9263094	.9564766	1.045504	.9684601
	16	.9026732	.9396132	1.064268	.9606859
	20	.8793852	.9216049	1.085064	.954189
	24	.8563823	.9024739	1.108065	.9489275
	28	.8336042	.8822432	1.133474	.9448688
	32	.8109925	.8609378	1.161524	.9419872
	36	.788491	.8385835	1.192487	.9402655
	40	.7660443	.8152073	1.226682	.9396927
	44	.7435978	.7908381	1.264481	.9402655

Thesis
W222333 Wallace
c.1 Investigation of computer vision methods for the building construction process.